Dissertation

Energy-Efficient Resource Allocation in Multi-Carrier Digital Subscriber Lines

ausgeführt zum Zwecke der Erlangung des akademischen Grades
eines Doktors der technischen Wissenschaften

eingereicht an der
Technischen Universität Wien
Fakultät für Elektrotechnik und Informationstechnik

von

Martin Wolkerstorfer

Wien, 16. Februar 2012

**TECHNISCHE
UNIVERSITÄT
WIEN**
Vienna University of Technology

Supervisor

Prof. Christoph Mecklenbräuker

Institute of Telecommunications
Vienna University of Technology, Austria

Examiner

Prof. Marc Moonen

Department of Electrical Engineering, ESAT/SCD
Katholieke Universiteit Leuven, Belgium

Kurzfassung

Digital subscriber lines (DSL) sind mit weltweit über 300 Millionen Nutzern eine der aktuell wichtigsten Breitbandtechnologien. Neben einem wachsenden Umweltbewusstsein ist es diese große Anzahl an bereits installierten Systemen, die Energieeffizienz zu einem wichtigen Gütemaß für DSL Systeme macht. Die Hauptbeeinträchtigung der Übertragungsqualität in gebräuchlichen DSL Systemen ist das Übersprechen im Kabelbündel. Die Koordinierung der Sendeleistungsspektra mehrerer Systeme erlaubt es, die nachteiligen Auswirkungen des Übersprechens abzuschwächen und die Übertragungseffizienz und Verbindungsstabilität zu verbessern. Das entsprechende mehrdimensionale "Rucksackproblem" ist jedoch nur für Zugangsnetze mit wenigen aktiven DSL Nutzern optimal lösbar. Das Ziel dieser Arbeit ist daher der Entwurf von robusten und energieeffizienten Algorithmen zum dynamischen Spektrum-Management (DSM) mit niedriger Komplexität, welche auf große DSL Netzwerke anwendbar sind.

Die vorgestellte Entwurfsstrategie zielt auf eine Sendeleistungsminimierung in Mehrträgerübertragungssystemen mit mehreren Nutzern ab, wobei eine Lagrange-Relaxierung des Optimierungsproblems zur Anwendung kommt. Die Wahl der Zielfunktion resultiert aus vergleichenden Simulationen mit einem nichtlinearen Modell des Leistungsverbrauchs eines Leitungstreibers, welcher einen Großteil der Gesamtleistung eines DSL Sendeempfängers darstellt. Es wird gezeigt, dass die verwendete Relaxierung eine "Dualitätslücke" in Bezug auf die Zielfunktion mit sich bringt, die jedoch in praxisrelevanten Szenarien vernachlässigbar ist. Um eine enge Schranke für die Performanz von DSM Algorithmen zu berechnen wird ein neuer Ansatz für dual-optimales DSM hergeleitet, der die Menge der derzeit lösbaren DSM Probleme erweitert. Für noch größere Netzwerke wird ein neuartiger und vielseitig anwendbarer DSM Algorithmus vorgeschlagen, welcher die Anwendung von effizienten Heuristiken erlaubt.

All diese DSM Verfahren erreichen Energieeffizienz durch die Reduktion der Sendeleistung bei gleichzeitiger Einhaltung vereinbarter, minimaler Bitraten. Von Netzbetreibern wird jedoch neben Energieeffizienz auch Serviceabdeckung nachgefragt. Darauf Rücksicht nehmend wird ein systematischer Ansatz vorgestellt, um zulässige minimale Bitraten zu berechnen. Dadurch lässt sich der Überschuss an Bitrate über das dem Service entsprechende Maß hinaus reduzieren und die damit einhergehende Vergeudung von Sendeleistung vermeiden. Nichtsdestotrotz werden energieeffiziente DSM Lösungen nur dann in der Praxis akzeptiert, wenn die nachgefragte Servicequalität auch unter Schwankungen von elektromagnetischen Störungen

und möglichen Fehlern im Optimierungsmodell bereitgestellt werden kann. Des Weiteren werden deshalb auf DSM basierende Techniken zur Verbindungsstabilisierung auf ihre Energieeffizienz untersucht. Es zeigt sich, dass der berechnete, damit verbundene zusätzliche Leistungsverbrauch gering ist im Vergleich zum Energieeinsparungspotential welches marktübliche DSL Technologien bergen.

Abstract

Digital subscriber lines (DSL) are one of the major current broadband access technologies, with a penetration of over 300 million DSL subscribers world-wide. Besides an increasing environmental awareness, it is this large number of deployed systems which has let energy-efficiency become an important operational metric for DSL. The major performance limiting impairment in current DSL systems is the interference among DSL lines. Using power spectral coordination among the lines we can mitigate the detrimental effects of interference and improve their efficiency and stability. The corresponding multi-dimensional nonlinear Knapsack problem however becomes intractable to be solved optimally for all but access networks with a few lines. The aim of this thesis is the design of novel robust and energy-efficient dynamic spectrum management (DSM) algorithms with low complexity for large DSL networks.

The presented design approach targets the transmit-power minimization and is based on a Lagrange dual relaxation of the multi-user and multi-carrier constrained optimization problem. The choice of objective is supported by simulations comparing it to a nonlinear model of the line-driver power consumption, a major contributor to the power budget of a DSL transceiver. The applied relaxation is shown to incur a gap in objective value compared to the original target function, which is however found to be negligible in practically relevant scenarios. In order to obtain a tight performance bound for DSM a new approach for dual optimal DSM is derived which enlarges the current set of DSM problems that can be solved optimally in practice. For even larger problems a novel versatile and scalable DSM framework is proposed which allows for the application of low-complexity heuristics.

All these DSM schemes are energy-efficient by reducing the transmit power while achieving an agreed target-rate. Recognizing the network operators' demand for both, energy-efficiency and service coverage, a systematic approach to find feasible target-rates is presented. This allows to lower the excess in rate beyond the targeted service and therefore to avoid the needless waste of transmit power. However, energy-efficient DSM solutions will only be accepted in practice if a demanded quality of service can be guaranteed under fluctuations in electromagnetic disturbances and possible errors in the optimization model. Hence, low-complexity DSM-based stabilization techniques are investigated in terms of their energy-efficiency, and found to expend little extra power compared to the savings achievable by currently available DSL technologies.

Acknowledgement

Contents

Contents

List of Figures

List of Tables

Abbreviations

Abbreviation	Description
ADC	Analog-to-Digital Conversion
ADSL	Asymmetric Digital Subscriber Line
AN	Artificial Noise
BFB	Best-First Branch-and-Bound
BnB	Branch-and-Bound
CHET	Combination Heuristic for Time-Sharing
CM	Coverage Maximization
CO	Central Office
CPU	Central Processing Unit
DAC	Digital-to-Analog Conversion
DBL	Discrete Bit-Loading
DCP	Difference-of-Convex-Functions Programming
DFB	Depth-First Branch-and-Bound
DMT	Discrete Multi-Tone Modulation
DSL	Digital Subscriber Line
DSLAM	DSL Access Multiplexer
DSM	Dynamic Spectrum Management
FDD	Frequency-Division Duplexing
FDMA	Frequency-Division Multiple Access
FEQ	Frequency Domain Equalizer
FEXT	Far-End Crosstalk
FFT	Fast Fourier Transformation

Table 0.1: Abbreviations A-F used in this thesis.

Abbreviation	Description
GP	Geometric Programming
Grasp	Greedy Randomized Search Procedure
ICT	Information and Communication Technology
IFFT	Inverse Fast Fourier Transformation
ISB	Iterative Spectrum Balancing
ISDN	Integrated Services Digital Network
ISI	Inter-Symbol Interference
JOGO	Joint Greedy Optimization
LD	Line-Driver
LDP	Line-Driver Power
LP	Linear Program
LPM	Low-Power Mode
LS (rLS)	(Randomized) Local Search
MDBL	Multi-user DBL
MEGS	Mixed Exhaustive and Greedy Search
NDW-DSM	Nonlinear Dantzig-Wolfe based DSM
NEXT	Near-End Crosstalk
NP	Non-Deterministic Polynomial-Time
OFDMA	Orthogonal FDMA
P/S	Parallel-to-Serial Conversion
PSD	Power Spectral Density
QAM	Quadrature Amplitude Modulation
RM	(Sum-)Rate Maximization
RSB	Regular-Splitting-based BnB
SDP	Semidefinite Programming
SEGO (rSEGO)	(Randomized) Sequential Greedy Optimization
SINR	Signal-to-Interference-plus-Noise Ratio
SMC	Spectrum Management Center
SNR	Signal-to-Noise Ratio
SOC	Second-Order Cone
S/P	Serial-to-Parallel Conversion
SSM	Static Spectrum Management
SSR	Search-Space Reduction
TP	Transmit Power
VDSL	Very High Speed DSL
VN	Virtual Noise
VoIP	Voice-over-Internet-Protocol
WS-LS	Warm-Start Local Search

Table 0.2: Abbreviations G-W used in this thesis.

Chapter 1 Introduction

1.1 Motivation and State-of-the-Art

Energy-efficiency is naturally an important design criterion for battery-powered communication systems. However, the growing pressure on the information and communication technology (ICT) sector to reduce its energy consumption has made researchers, standardization bodies, companies, and governmental bodies also think about the energy-efficiency of wired communication systems [48, 51, 83–85, 89]. For example, the European "Code of Conduct on Energy Consumption of Broadband Equipment", a voluntary governmental initiative existing since July 2006, has set its goal to halve the expected electricity consumption of broadband equipment by 2015 [48]. In [64] the share of the fixed broadband access in the telco's energy consumption for 2020 is estimated at around $14\,\%$. Herein we focus on the most widely deployed fixed broadband access technology today [170], namely digital subscriber lines (DSL).

Approaches for reducing the power consumption of DSL equipment can be classified into three categories [20]: the optimization of hardware components (or also network topology), dynamic rate adaptation (or spectral optimization, respectively), and low-power operation modes. Up to the beginning of the work on this thesis only little effort has been made to systematically optimize the energy-efficiency in DSL. A noteworthy exception is the introduction of low-power modes (LPM) in the asymmetric DSL (ADSL) 2 standard [86], a technique facing various concerns related to the instability it introduces in the network. Since then various proposals were made on how to save energy in DSL, for instance by the design of energy-efficient hardware modules [134, 137, 147, 185], by using LPMs and enabling its stable usage [18, 58, 69, 95, 189], by the dimensioning and energy-efficient operation of the network processor [76], by wireless traffic aggregation at the user side and efficient line-card usage at the network side [61], or by the deployment of street cabinets [15, 62, 74]. The latter allows to reduce the transmit power by shortening the cable length, and to reduce cooling requirements [16, 17] by reducing the number of installed line cards. In this thesis we restrict ourselves to a specific means of tackling the energy minimization problem in DSL, namely by judiciously controlling the transmit power, also known as *dynamic spectrum management* (DSM) [149]. The complexity of the optimal DSM problem [33] has triggered numerous studies on approximate and fast DSM algorithms. However, multi-user DSM in interference-limited multi-carrier communication systems such as DSL remains a challenging

problem, especially so in networks with many users and with respect to optimality guarantees.

Besides our initial work in [186], the idea of reducing the transmit power by DSM was also followed by others, for example in [67, 164]. In [65, 68] DSM is performed with the objective to optimize a functional model of the line-driver (LD) power consumption, which constitutes a large part of the power budget in today's DSL transceivers. From power consumption measurements of today's state-of-the-art very high speed DSL (VDSL) 2 chipsets we know that the LD accounts for up to 60 % [186]. A reduction in transmit power was shown to result in additional energy savings by reducing the power consumption of the LD [67, 186]. Thus, by DSM we can significantly lower the DSL *system* power consumption as a whole. However, an open question remains which information needs to be exchanged between protocol layers (e.g., on the used hardware) to design energy-efficient DSM algorithms.

In the form considered in this thesis, DSM necessitates the knowledge of the magnitudes of the channel transfer coefficients, and the control of the transmitted power levels. Implementations of DSM in real systems have been described in [30, 117, 154, 168, 175]. While in [168] iterative water-filling was applied in ADSL, in [154] the standardized power back-off technique was used in VDSL2 systems. Both DSM realizations are applicable without explicit knowledge of the crosstalk channel. Differently, two DSM heuristics were applied to ADSL2 systems in [30, 117] which make use of estimated crosstalk coefficients. An open question remains how crosstalk channel identification errors observed in practice [117], [92, Ch. 3] can be incorporated in DSM algorithms and how they impact the energy-efficiency of the DSM technique.

1.2 Models and Notation

1.2.1 DSL System Model

We consider a modern DSL transceiver as illustrated in Figure 1.1, cf. [59, 60] for details on the following system model. Multi-carrier modulation is implemented using the discrete multi-tone (DMT) technique, where an inverse fast Fourier transform (IFFT) and FFT pair is used to effectively split the spectrum into C independent narrowband subchannels ("subcarriers"). The input to the IFFT is generated by allocating a block of data bits to the C subcarriers and quadrature amplitude modulation (QAM), scaled to control the desired digital power of the symbol (commonly referred to as "power-loading"). The process of mapping an integer number of bits to subcarriers is referred to as "bit-loading". Integer modulation constellations allow for simpler decoder implementations [59]. Bit-loading is based on the estimated channel frequency response and spectral noise power on each subcarrier and accomplished during the initialization phase of a line. The specific IFFT gives real output symbols which are extended to combat inter-symbol interference (ISI) and
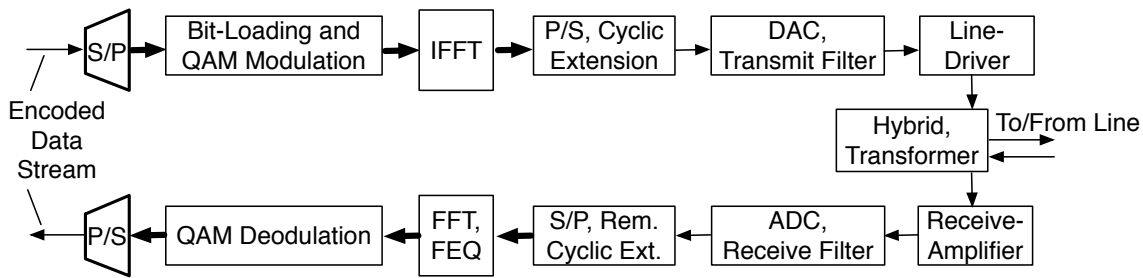
Figure 1.1: Schematic of a DSL transceiver, highlighting various signal processing blocks in the transmission and reception path, such as the serial-to-parallel (S/P) and parallel-to-serial (P/S) conversion, the quadrature amplitude modulation (QAM), the (inverse) fast Fourier transformation ((I)FFT), the digital-to-analog conversion (DAC) and analog-to-digital conversion (ADC), and the frequency domain equalizer (FEQ).

to enable synchronization between the transmitter and the receiver. The signal is in the following digitally filtered, digital-to-analog (DAC) converted, and runs through an analog filter to match signal requirements such as a regulatory power spectral mask. The line-driver serves to amplify the signal and to match the characteristic impedance to the line. Finally, the hybrid circuit enables the simultaneous transmission and reception of signals in the frequency-division duplexing (FDD) system by attenuating the leakage of the stronger transmit signal into the receiver path. The equivalent reverse operations are performed at the receiver side to recover the received, encoded data stream.

In this thesis we study the problem of multi-user bit-loading and power-loading, which we refer to as "dynamic spectrum management" (DSM) [40, 125, 149]. More precisely, our definition of DSM corresponds to what is known as "spectrum balancing" [152], also referred to as DSM with "level 2" coordination, cf. [98] for an overview. DSM is a part of dynamic line/bundle management [116, 159] which refers more generally to any kind of physical-layer parameter adaptation for rate and quality improvement, including for example also changes of rate and code settings.

1.2.2 DSL Network and Channel Model

Simple network models are shown in Figure 1.2, consisting of U lines deployed in a joint cable binder. The transmission direction from the central office (CO) to the subscribers is referred to as "downstream", while that in the reverse direction is referred to as "upstream". By electromagnetic radiation the signal on one line couples into the other lines. While the copper wires are in fact twisted to mitigate this effect, imperfect twisting leads to a remaining interference, commonly referred to as "crosstalk". Assuming perfectly synchronized DMT modulation at each modem and the use of FDD, the near-end crosstalk (NEXT) is negligible, that is the interference
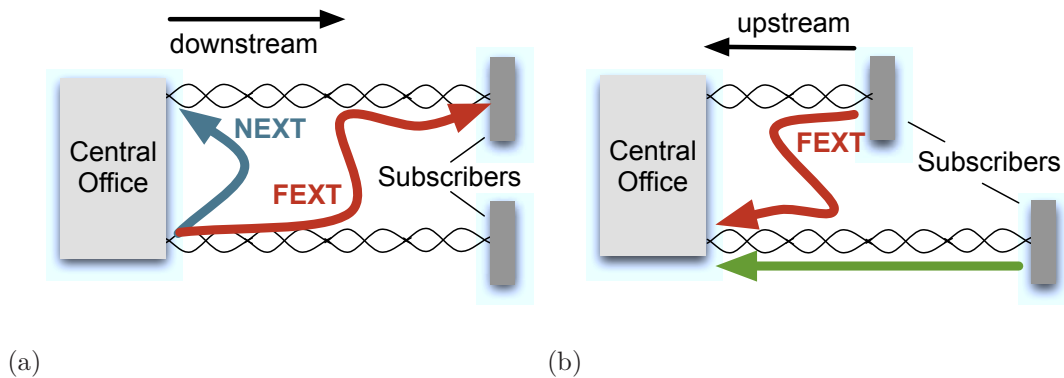
Figure 1.2: Schematic crosstalk scenarios; (a) Near-end crosstalk (NEXT) and far-end crosstalk (FEXT); (b) Upstream near-far scenario.

between the collocated transmitter and receiver, cf. Figure 1.2(a). Further assuming a sufficient length of the cyclic prefix allows to neglect ISI and to model the subcarriers as orthogonal subchannels. Hence, we effectively face a far-end crosstalk (FEXT) limited system due to the typically comparably low background-noise levels, where FEXT refers to the crosstalk between transmitters and receivers on opposing ends of the cable, cf. Figure 1.2(a). We assume a central unit, e.g. a spectrum management center (SMC) at the collocated end of the cable bundle, has full knowledge of the magnitudes of the crosstalk couplings. The most critical FEXT scenario occurs in so-called "near-far" situations – that is the case when the intended received signal is strongly attenuated over a long loop, while the received crosstalk noise is relatively strong as the distance between the receiver and the disturbing transmitter is relatively short, cf. Figure 1.2(b). Near-far situations occur in practice for example in upstream transmission when the loop-lengths vary widely among subscribers, and mixed central office and cabinet deployments [40] where the crosstalk victim and disturber are not collocated on either end of their lines.

An important parameter for the simulation of DSM-enabled DSL systems is the chosen FEXT model as it influences the coupling among users, the achievable performance gain by DSM compared to static (i.e., network-oblivious) power-allocation rules, as well as the complexity of optimally solving the DSM problem as we shall see in Chapter 3. We mostly rely on the widely used and reproducible and empirical 99 % worst-case channel[1] model [59]. Denoting the channel coefficient from user i to user u by G_{ui}^c, the transfer coefficient $H_{ui}^c = |G_{ui}^c|^2$ on subcarrier c is modeled as

$$H_{ui}^c = K_{FEXT} \cdot f_c^2 \cdot L \cdot H_{uu}^c, \tag{1.1}$$

---

[1]This means that in no more than 1 % of the taken channel measurements the FEXT transfer function exceeded the given model.
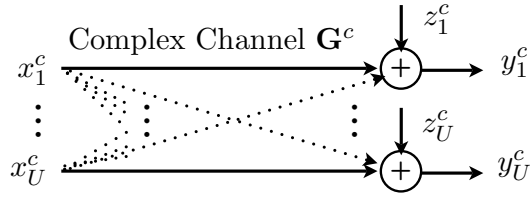
Figure 1.3: Gaussian interference channel on subcarrier c due to electromagnetic coupling among U copper lines.

where K_{FEXT} is a constant with suggested value for Europe [59] of $K_{FEXT} = 10^{-4.5}$, f_c [MHz] is the center frequency of the c'th subcarrier, and L [km] is the loop-length over which the interfering lines are coupled. The value $H_{uu}^c = |G_{uu}^c|^2$ is the transfer coefficient of the direct channel at frequency f_c, commonly given by a two-port transmission line model [59], with parameters following the "TP100" cable in [50]. In general the signal attenuation over the direct channel is larger at higher frequencies. The channel models in this thesis were calculated using the xDSL simulator publicly available in [126]. Our intention behind the usage of the worst-case FEXT model is to construct challenging interference scenarios for the design and comparison of DSM algorithms. Simulation results which show the gain in power consumption by DSM should therefore be carefully interpreted, although in [115] it has been reported that in Europe even higher crosstalk levels than expected from the worst-case model are often observed in the field. However, note that the power reduction by power-minimizing DSM under this model compared to full-power transmission is indeed an underestimate of the potential savings by DSM, as larger crosstalk couplings imply an increase in the power consumption at the power-minimizing DSM solution. Hence, the worst-case model is used throughout this thesis if not mentioned otherwise. More precisely, in Sections 2.3.4, 3.3.1, 3.3.4, and 3.3.5 we use measured coupling data among 24 lines taken from [153], for instance to demonstrate the optimal solvability of the DSM problem under low crosstalk couplings. This has however the drawback of giving non-reproducible results. Both, reproducible as well as measurement-based statistical FEXT models have been reported in [115,150,191] and could alternatively be used in Monte-Carlo simulations to quantify the average performance gain by DSM. Differently, we will at various places show average results under uniformly sampled loop-lengths L for each user, but using the worst-case model above.

Under the made assumptions the equivalent baseband model for subcarrier $c \in \mathcal{C} = \{1, \ldots, C\}$ is in the form of an interference channel [43], cf. Figure 1.3, where the received symbol of user $u \in \mathcal{U} = \{1, \ldots, U\}$ is given as

$$y_u^c = G_{uu}^c x_u^c + \sum_{i \in \mathcal{U} \backslash u} G_{ui}^c x_i^c + z_u^c, \tag{1.2}$$

where the transmitted symbols are denoted as $x_u^c \sim \mathcal{N}(0, p_u^c)$, $z_u^c \sim \mathcal{N}(0, N_u^c)$, is

19

the additive white Gaussian background noise, and $\mathcal{N}(\mu, \sigma^2)$ represents a circular symmetric complex normal distribution with mean μ and variance σ^2.

1.2.3 DSL Optimization Model

The DSL channel can be assumed to be only slowly time-varying, the main reason being temperature changes. Hence, it is reasonable to optimize all users' bit-allocations ("bit-loading") and corresponding power-allocations ("power-loading") based on the estimated transfer coefficients, cf. [106] for a practical estimation method. With the exception of Chapter 6 we will assume that these coefficients are perfectly known. Online reconfiguration methods [59, Sec. 7.5] adapt the bit-allocation and the power-allocation continuously and therefore complement the "offline" DSM problem considered in this thesis. In [40] it is stated that the DSM parameter recommendations might change "daily or even more often", which together with the slow variability of the channel supports the practical feasibility of using sophisticated offline optimizing DSM schemes.

In our DSL model the modems regard crosstalk solely as noise, which, for a sufficiently high number of users, can be well approximated by a Gaussian distribution, a common approximation in spectrum management applications [152, Ch. 11]. Hence, the achievable rate per DMT-symbol for user $u \in \mathcal{U}$ on subcarrier $c \in \mathcal{C}$ assuming two-dimensional signal constellations can be approximated in practice [59] by

$$r_u^c(\mathbf{p}^c) \doteq \log_2\Big(1 + \frac{H_{uu}^c p_u^c}{\Gamma(\sum_{i \in \mathcal{U} \backslash u} H_{ui}^c p_i^c \; + \; N_u^c)}\Big), \tag{1.3}$$

where $\mathbf{p}^c = [p_1^c, \ldots, p_U^c]^T$, p_u^c is the power spectral density (PSD) of the transmitted signal, Γ is the signal-to-noise ratio (SNR) gap which allows to effectively approximate the performance loss compared to the theoretical capacity [43], N_u^c is the total receiver noise spectral density, and H_{uu}^c and H_{ui}^c are the squared magnitudes of the channel transfer coefficient of user u and from user i to user u, respectively, on subcarrier c. We will write the vector of all users' rates as $\mathbf{r}^c(\mathbf{p}^c) = [r_1^c(\mathbf{p}^c), \ldots, r_U^c(\mathbf{p}^c)]^T$, and use $\mathbf{p}^c(\mathbf{r}^c)$ to denote the unique [194] power-allocation resulting in the rate vector $\mathbf{r}^c$. By adequate reformulation of the equation in (1.3) the computation of $\mathbf{p}^c(\mathbf{r}^c)$ can be written as the single-carrier power control problem [37, 54]

$$\underset{\mathbf{p}^c \succeq \mathbf{0}}{\text{minimize}} \quad \sum_{u \in \mathcal{U}} p_u^c \tag{1.4a}$$

$$\text{subject to} \quad (\mathbf{I} - \mathbf{F}^c) \, \mathbf{p}^c \succeq \mathbf{n}^c, \tag{1.4b}$$

where $\mathbf{F}^c \in \mathbb{R}^{U \times U}$, with elements $F_{uv}^c = \Gamma \gamma_u^c \frac{H_{uv}^c}{H_{uu}^c}, c \in \mathcal{C}, \; u \in \mathcal{U}, v \in \mathcal{U}$, where $\mathbf{n}^c = [\Gamma \gamma_1^c \frac{N_1^c}{H_{11}^c}, \ldots, \Gamma \gamma_U^c \frac{N_U^c}{H_{UU}^c}]^T$, $\mathbb{R}$ denotes the set of real numbers, and where $\gamma_u^c = (2^{r_u^c} - 1)$ is the targeted signal-to-interference-plus-noise ratio (SINR) of user u on subcarrier

c. While the SINR is constant in this single-carrier problem, we will later see how the spectrum balancing algorithm optimizes and determines the values r_u^c and hence γ_u^c, $c \in \mathcal{C}, u \in \mathcal{U}$. The SINR-constraints in the linear program (LP) in (1.4) keep the objective from attaining the minimum it would obtain when neglecting the SINR-constraints, which due to $\mathbf{p}^c \succeq \mathbf{0}$ is 0. Hence, at optimum the constraints in (1.4b) hold with equality. Furthermore, the unique optimal solution in (1.4) is element-wise lower or equal than any other feasible solution [103]. From this it follows that adding linear weights in the objective or a maximum per-user power restriction does not change this optimal solution. Provided the problem in (1.4) is feasible, the optimal solution $\mathbf{p}^c(\mathbf{r}^c)$ of (1.4) is given as the solution of the linear matrix equality (cf. also [22, Ex. 4.8])

$$(\mathbf{I} - \mathbf{F}^c)\,\mathbf{p}^c = \mathbf{n}^c. \tag{1.5}$$

It is well known [37,54] that, assuming *nonzero* background noise, crosstalk couplings and SINR-targets, a solution to the power control problem in (1.4) exists iff

$$\rho\left(\mathbf{F}^c\right) < 1, \tag{1.6}$$

where $\rho(\cdot)$ denotes the spectral radius[2] or Perron-Frobenius eigenvalue. To model constraints on the modulation's constellation size and regulatory PSD mask limitations we use the set of feasible per-subcarrier PSDs on subcarrier $c \in \mathcal{C}$,

$$\mathcal{Q}^c = \{\mathbf{p}^c | r_u^c\left(\mathbf{p}^c\right) \in \mathcal{B},\ 0 \le p_u^c \le \hat{p}_u^c, \forall u \in \mathcal{U}\}, \tag{1.7}$$

where $\hat{p}_u^c, c \in \mathcal{C}, u \in \mathcal{U}$, indicates the PSD mask, and $\mathcal{B} = \{0, \theta, 2\,\theta, \ldots, \hat{\theta}\}$ is the set of discrete per-subcarrier bit-allocations with rate-steps of size θ and a bit-cap $\hat{\theta}$.

1.2.4 Generic Multi-Carrier Optimization Problems

At the center of this thesis is a generic multi-user DSM problem in the form of a multi-dimensional nonlinear Knapsack problem [123], formulated as

$$P_{(\mathbf{R},\hat{\mathbf{P}})}^* = \underset{\mathbf{p}^c \in \mathcal{Q}^c, c \in \mathcal{C}}{\text{minimize}} \quad \sum_{c \in \mathcal{C}} f(\mathbf{p}^c, \hat{\mathbf{w}}, \check{\mathbf{w}}) \tag{1.8a}$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} r_u^c\left(\mathbf{p}^c\right) \ \ge \ R_u, \quad \forall u \in \mathcal{U}, \tag{1.8b}$$

$$\sum_{c \in \mathcal{C}} p_u^c \ \le \ \hat{P}_u, \quad \forall u \in \mathcal{U}, \tag{1.8c}$$

---

[2]The spectral radius of a matrix is the largest magnitude of any of its eigenvalues [2].

where the objective is a weighted-sum of sum-powers and sum-rates defined by

$$f^c(\mathbf{p}^c, \hat{\mathbf{w}}, \check{\mathbf{w}}) = \hat{\mathbf{w}}^T \mathbf{p}^c - \check{\mathbf{w}}^T \mathbf{r}^c(\mathbf{p}^c), \ c \in \mathcal{C}, \tag{1.9}$$

where $\hat{\mathbf{w}}, \check{\mathbf{w}} \in \mathbb{R}_+^U$ are weights, where $\mathbf{R} \in \mathbb{R}_+^U$ are the target-rates in [bits/DMT-symbol], and where $\hat{\mathbf{P}} \in \mathbb{R}_+^U$ are the maximum sum-powers supported by the line-driver. Note that this formulation allows to consider sum-power minimizing users as well as sum-rate maximizing users in a single optimization problem by an adequate setting of the weights $\hat{\mathbf{w}}$ and $\check{\mathbf{w}}$, which results in a trade-off between the two objectives. This captures for example the problem in [158] where resource allocation under inhomogeneous traffic is studied. While a rate maximization problem is considered for best-effort users, delay-constrained users are incorporated at the same time under a minimum-rate constraint. In [166] various DSM formulations similar to that in (1.8) are presented which target a fair power reduction in DSL networks. Intuitively, a solution is considered fair in [166] if a rate reduction compared to an operation point on the rate-region boundary is accompanied by an adequate power reduction. While in the objective in (1.9) we only consider the *transmit* power, in Section 2.3 we will discuss its relation to the system power consumption.

Efficiently computable special cases of (1.8) include the single-user case as shown in Section 2.1.2, and the analytically solvable single-carrier power-control problem [37] in (1.4). Note that $r_u^c(\mathbf{p}^c)$ in (1.3) is a quasi-convex function[3], resulting in general in a non-convex objective in (1.8a) in the variables $\mathbf{p}^c, c \in \mathcal{C}$. This, besides the discrete bit-loading, can be seen as the origin of the non-convexity of the problem in (1.8) which persists even in the continuous bit-loading case.

For later reference we define the Lagrange dual problem to (1.8) as

$$D^*_{(\mathbf{R}, \hat{\mathbf{P}})} = \underset{\boldsymbol{\lambda} \succeq \mathbf{0}, \boldsymbol{\nu} \succeq \mathbf{0}}{\text{maximize}} \quad q(\boldsymbol{\lambda}, \boldsymbol{\nu}), \tag{1.10}$$

where the (partial) dual function [12, 22, 124] is defined as

$$q(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \sum_{c \in \mathcal{C}} q^c(\boldsymbol{\lambda}, \boldsymbol{\nu}) + \boldsymbol{\lambda}^T \mathbf{R} - \boldsymbol{\nu}^T \hat{\mathbf{P}}, \tag{1.11}$$

where $\boldsymbol{\lambda}, \boldsymbol{\nu} \in \mathbb{R}^U$ are the Lagrange multipliers associated with the constraints in (1.8b) and (1.8c), respectively, and where

$$q^c(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \min_{\mathbf{p}^c \in \mathcal{Q}^c} \left\{ f^c(\mathbf{p}^c, \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda}) \right\}, \quad \forall c \in \mathcal{C}. \tag{1.12}$$

The optimal dual objective $D^*_{(\mathbf{R}, \hat{\mathbf{P}})}$ lower-bounds the optimal primal objective $P^*_{(\mathbf{R}, \hat{\mathbf{P}})}$ in (1.8a) as given by the weak-duality inequality [12] $D^*_{(\mathbf{R}, \hat{\mathbf{P}})} \leq P^*_{(\mathbf{R}, \hat{\mathbf{P}})}$. The difference

---

[3]This follows from quasi-convexity of the logarithm's argument in (1.3) [22, Ex. 3.32] and composition with the (nondecreasing) logarithm [22, Sec. 3.4.4].

between the two objectives is referred to as the "duality-gap" [12], given by

$$\zeta = P^*_{(\mathbf{R},\hat{\mathbf{P}})} - D^*_{(\mathbf{R},\hat{\mathbf{P}})} \geq 0. \tag{1.13}$$

We refer to Section 2.1.1 for a derivation of (1.13) and more details on Lagrange relaxation.

1.3 Outline and Contributions

This thesis is based on work which is partly published in the following peer-reviewed papers:

[W1] M. Wolkerstorfer, J. Jaldén, and T. Nordström, "Column Generation for Discrete-Rate Multi-User and Multi-Carrier Power Control", *accepted for publication in IEEE Transactions on Communications*, January 2012.

[W2] M. Wolkerstorfer, T. Nordström, B. Krasniqi, M. Wrulich, and Ch. Mecklenbräuker, "OFDM/OFDMA subcarrier allocation", In *Cross Layer Designs in WLAN Systems*, Vol. 1, N. Zorba, Ch. Skianis, Ch. Verikoukis, Eds., Troubador Publishing Ltd, 2011, p. 177–216, chapter 6.

[W3] M. Wolkerstorfer, B.D. Putra, T. Nordström, and S. Trautmann, "Modeling and Optimization of the Line-Driver Power Consumption in xDSL Systems", *submitted to EURASIP Journal on Advances in Signal Processing*, February 2012.

[W4] M. Wolkerstorfer, D. Statovci, and T. Nordström, "Energy-Saving by Low-Power Modes in ADSL2", *accepted for publication in Elsevier Computer Networks, Special Issue on Green Communication Networks*, January 2012.

[W5] M. Wolkerstorfer, D. Statovci, and T. Nordström, "Dynamic Spectrum Management for Energy-Efficient Transmission in DSL", In *Proc. International Conference on Communications Systems (ICCS)*, Guangzhou, China, November 19-21, 2008.

[W6] M. Wolkerstorfer, D. Statovci, and T. Nordström, "Duality-gap bounds for multi-carrier systems and their application to periodic scheduling", In *Proc. IEEE International Conference on Communications (ICC)*, Cape Town, South Africa, May 23-27, 2010.

[W7] M. Wolkerstorfer, J. Jaldén, and T. Nordström, "Low-Complexity Optimal Discrete-Rate Spectrum Balancing in Digital Subscriber Lines", *submitted to Elsevier Signal Processing*, October 2011.

[W8] M. Wolkerstorfer and T. Nordström, "Coverage Optimization in DSL Networks by Low-Complexity Discrete Spectrum Balancing", In *Proc. IEEE Global Communications Conference (Globecom)*, Houston, Texas, USA, December 5-9, 2011.

[W9] M. Wolkerstorfer and T. Nordström, "Heuristics for Discrete Power Control - A Case-Study in Multi-Carrier DSL Networks", In *Proc. ALIO/EURO Workshop on Applied Combinatorial Optimization*, Porto, Portugal, May 4-6, 2011.

[W10] M. Wolkerstorfer, D. Statovci, and T. Nordström, "Robust spectrum management for DMT-based systems", *IEEE Transactions on Signal Processing*, 58(6): 3238–3250, June 2010.

[W11] M. Wolkerstorfer, D. Statovci, and T. Nordström, "Enabling Greener DSL Access Networks by their Stabilization with Artificial Noise and SNR Margin", *submitted to Springer Cluster Computing, Special Issue on Optimization Issues in Energy Efficient Distributed Systems*, November 2011.

Other authored/co-authored publications which are related to this thesis are listed in the following:

- G. Marrocco, M. Wolkerstorfer, T. Nordström, and D. Statovci, "Energy-Efficient DSL using Vectoring", In *Proc. IEEE Global Communications Conference (Globecom)*, Houston, Texas, USA, December 5-9, 2011.

- B. Krasniqi, M. Wolkerstorfer, C. Mehlführer, C. Mecklenbräuker "Sum-Rate Maximization for Multiple Users in Partial Frequency Reuse Cellular Networks", In *Proc. IEEE Broadband Wireless Access Workshop*, Miami, Florida, USA, December 6, 2010.

- M. Wolkerstorfer, T. Nordström, and D. Statovci, "Delay-constrained scheduling for interference-limited multi-carrier systems", In *Proc. IEEE International Workshop on Cross-Layer Design (IWCLD)*, Palma de Mallorca, Spain, June 11-12, 2009.

- T. Nordström, D. Statovci, and M. Wolkerstorfer, "Energy Efficient Power Back-Off Management for VDSL2 Transmission", In *Proc. European Signal Processing Conference (EUSIPCO)*, Glasgow, UK, August 24-28, 2009.

In the following we summarize the content of each chapter in more detail:

The first part of this thesis in **Chapter 2** summarizes our motivation for the concept of transmit-power efficient DSM in DSL, derived from various view-points. We provide a mathematical problem formulation, survey the common algorithmic

DSM approach based on Lagrange-dual relaxation, analyze its complexity, derive its relation to the time-sharing relaxation, and study the duality-gap in our non-convex DSM problem. The latter empirically justifies the application of Lagrange-dual relaxation for DSM. While the transmit power constitutes only a minor fraction of the transceiver's total power consumption, we will show that DSM significantly reduces the consumption beyond the transmit power. Also, empirical evidence shows that knowledge of the line-driver hardware does not significantly increase the possible energy savings by transmit-power efficient DSM, which motivates the DSM algorithms developed in Part II. Yet another motivation for our DSM framework presented in this chapter is the low-power state in DSL, that is an operation mode where the modem operates at a reduced transmit power and rate. However, this mode leads to more severe fluctuations in crosstalk noise, and hence necessitates active, DSM-based stabilization techniques. One such approach will be described in more detail in Chapter 7. Chapter 2 is to a large extent based on [W1–W6].

The second part of this thesis is concerned with the development of efficient algorithms for the DSM problem in large DSL networks with versatile optimization objectives.

Chapter 3 describes a novel algorithm for optimally solving the combinatorial per-subcarrier power control problems. Previously proposed algorithms for its optimal solution are only applicable for networks with few users, while the suboptimality of less complex bit-loading algorithms has not been studied adequately so far. We deploy problem-specific branch-and-bound and search-space reduction methods which for the first time give a low-complexity guarantee of optimality in certain multi-user DSL networks of practical size. Simulation results show the dependency of our algorithm's complexity on the target-rate, and precisely quantify the suboptimality of multi-user bit-loading schemes in thousand ADSL2 scenarios under measured channel data. This chapter is based on [W7].

Chapter 4 summarizes a novel framework for discrete-rate spectrum balancing. More precisely, a column generation based DSM algorithm based on the time-sharing relaxation as well as a heuristic to recover primal feasible solutions is described. Low-complexity combinatorial rate and power-allocation heuristics for the single-carrier problem are proposed and their performance compared by simulation. Simulation results in randomly generated DSL network topologies are presented which show the improvements by our column generation scheme using the allocation heuristics in terms of sum-rate and sum-power compared to state-of-the-art DSM algorithms. This chapter is based on [W1], [W8], and [W9].

Chapter 5 describes a novel problem formulation for DSM, targeting the maximization of service coverage. A low-complexity heuristic is presented which is

based on the column generation framwork of Chapter 4 and a convex relaxation at the initial stage. Two relaxation approaches are compared to the global optimum in small-sized networks by simulation. For large networks we compare the coverage maximization approach using heuristics for the per-subcarrier problems to sum-rate maximizing schemes in terms of coverage and power consumption. Corresponding simulation results demonstrate an average gain in service coverage of more than 13 % and an average transmit-power reduction by 37 % compared to sum-rate maximizing DSM. This chapter is partly based on [W8].

In the third part of this thesis we study two cases where the actual DSL environment may differ from that when the power-allocation was optimized offline.

In **Chapter 6** we propose the concept of a crosstalk margin which conservatively captures our uncertainty in the measured crosstalk couplings. We exemplify the concept through two computationally advantageous uncertainty models and demonstrate the trade-off between the rate / sum-power performance and robustness. Simulation results confirm that the transmit power can be reduced while providing robustness with respect to crosstalk uncertainty. This chapter is partly based on [W10].

Chapter 7 deals with the stabilization of DSL networks under varying crosstalk noise. The injection of artificial noise (AN) at the transmitter has been proposed as a solution for current DSL systems without the need for further standardization. We develop the idea of setting the frequency selective AN power jointly with the SNR margin, based on a worst-case stabilization criterion. Furthermore, the AN is adjusted jointly with the transmit-power spectrum and the frequency flat SNR margins in order to compute a performance bound for AN-enabled networks. Simulation results confirm the strong dependency of the performance under AN on the selected SNR margin, which motivates their joint optimization for the actual network topology. This chapter is based on [W11].

Part I

Energy-Efficient Dynamic Spectrum Management (DSM)

Chapter 2 Lagrange Relaxation based Transmit-Power Optimization

We begin this chapter by introducing Lagrange relaxation based DSM in Section 2.1 and, differently to previous work, motivate it from various view-points. Its initial motivation [33] is the complexity reduction by decomposing the DSM problem into many per-subcarrier problems. Additionally to that we will a) show its equivalence to the time-sharing relaxation which opens the way for the improved DSM implementations proposed in Chapter 4, b) empirically demonstrate a small gap in objective value to the original problem in Section 2.2, concluding that DSM has sufficient capabilities for interference-avoidance in current DSL systems, c) motivate it by comparing the achieved line-driver (LD) power consumption to a direct optimization of the LD power in Section 2.3, d) motivate it by the substantial achieved power reduction in Section 2.3.4, and e) briefly motivate it as an enabling technique for the energy reductions promised by low-power modes in Section 2.4, on which we further elaborate in Chapter 7. At the core of all these motivational examples are the primal and Lagrange-dual DSM problems introduced in Section 1.2.4 and analyzed more closely in the following section.

The application of Lagrange relaxation in power spectrum shaping for improving the spectral compatibility between different DSL technologies has been proposed in [7]. In [33] Lagrange relaxation was applied to a combinatorial multi-user DSM problem in modern DSL systems, which triggered further research on optimal and/or low-complexity Lagrange-dual relaxation based DSM algorithms for DSL. We refer to [31, 163, 198] for examples of discrete-rate Lagrange relaxation based optimization algorithms, and to [82, 129, 161] for an overview of various continuous optimization schemes for the DSM problem.

2.1 Mathematical Problem Formulations for DSM

2.1.1 Background Information on Lagrange Relaxation

In this section we elaborate more on the initial primal and Lagrange-dual DSM problem formulations in (1.8) and (1.10), respectively. The separable function

$L(\mathbf{p}, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) = \sum_{c \in \mathcal{C}} f^c(\mathbf{p}^c, \hat{\mathbf{w}} + \boldsymbol{\nu}, \breve{\mathbf{w}} + \boldsymbol{\lambda}) + \boldsymbol{\lambda}^T \mathbf{R} - \boldsymbol{\nu}^T \hat{\mathbf{P}}$ which we minimize in (1.12) is referred to as "Lagrangian" in the literature [12], where $\mathbf{p} = [(\mathbf{p}^1)^T, \dots, (\mathbf{p}^C)^T]^T$. More precisely, in (1.12) the point-wise minimum is taken over a set of functions $f^c(\mathbf{p}^c, \hat{\mathbf{w}} + \boldsymbol{\nu}, \breve{\mathbf{w}} + \boldsymbol{\lambda})$ which are linear in the Lagrange multipliers $\boldsymbol{\nu}$ and $\boldsymbol{\lambda}$, respectively, and indexed by the power-allocations $\mathbf{p}^c \in \mathcal{Q}^c$. Therefore the dual function $q(\boldsymbol{\lambda}, \boldsymbol{\nu})$ in (1.11) is always concave [12], implying that the dual problem in (1.10) is a convex optimization problem. Denoting a pair of primal and dual solutions by $\mathbf{p}^*$ and $(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*)$, respectively, we have the relations

$$q(\boldsymbol{\lambda}, \boldsymbol{\nu}) \leq q(\boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) = \min_{\mathbf{p}^c \in \mathcal{Q}^c, c \in \mathcal{C}} \left\{ L(\mathbf{p}, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \right\} \tag{2.1a}$$

$$\leq L(\mathbf{p}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) \tag{2.1b}$$

$$\leq \sum_{c \in \mathcal{C}} f^c(\mathbf{p}^{c,*}, \hat{\mathbf{w}}, \breve{\mathbf{w}}) = P^*_{(\mathbf{R}, \hat{\mathbf{P}})}, \tag{2.1c}$$

where the inequality in (2.1c) follows from the feasibility of $\mathbf{p}^*$ for the original constraints in (1.8b) and (1.8c), respectively. From these relations it follows that the dual function $q(\boldsymbol{\lambda}, \boldsymbol{\nu})$ is a lower bound to the primal optimum $P^*_{(\mathbf{R}, \hat{\mathbf{P}})}$, and also the weak-duality inequality in (1.13). Moreover, in the special case where the solution $\mathbf{p}(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ obtained from the subproblems in (1.12) for non-negative (that is, feasible for the dual problem in (1.10)) variables $\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}$ is primal feasible in (1.8) and complementarity[1] holds, then $\mathbf{p}(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}})$ is also optimal for the primal problem in (1.8). This holds by (1.13) and

$$D^*_{(\mathbf{R}, \hat{\mathbf{P}})} = \underset{\boldsymbol{\lambda} \succeq \mathbf{0}, \boldsymbol{\nu} \succeq \mathbf{0}}{\text{maximize}} \, q(\boldsymbol{\lambda}, \boldsymbol{\nu}) \geq q(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) = L(\mathbf{p}(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}), \tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) \tag{2.2a}$$

$$= \sum_{c \in \mathcal{C}} f^c(\mathbf{p}^c(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}), \hat{\mathbf{w}}, \breve{\mathbf{w}}) \tag{2.2b}$$

$$\geq P^*_{(\mathbf{R}, \hat{\mathbf{P}})}, \tag{2.2c}$$

where (2.2b) holds due to the assumed complementarity. A consequence of a positive duality-gap ζ is that the dual function $q(\boldsymbol{\lambda}, \boldsymbol{\nu})$ in (1.11) is non-differentiable at every dual optimal solution [12, Ex. 6.1.1]. Furthermore, it is polyhedral[2] as a consequence of the finiteness of $\mathcal{Q}^c, c \in \mathcal{C}$, cf. [12, Sec. 6.1]. Both facts give reason for the application of non-differentiable optimization algorithms [12] to the dual problem in (1.10). We refer to [12] for more details on Lagrange-dual relaxation.

---

[1]Complementarity [12, Sec. 3.3.1] holds if $\sum_{c \in \mathcal{C}} r_u^c \left(\mathbf{p}^c(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) \right) = R_u$ if $\tilde{\lambda}_u > 0$, and $\sum_{c \in \mathcal{C}} p_u^c(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) = \hat{P}_u$ if $\tilde{\nu}_u > 0$, $\forall u \in \mathcal{U}$.

[2]That is, conv$(\breve{\mathcal{Q}})$ is a finitely constrained polyhedron [12, p. 719], where conv$(\cdot)$ denotes the convex-hull operation and $\breve{\mathcal{Q}}$ is defined in (2.5).

2.1.2 Time-Sharing Relaxation

As an alternative relaxation to the dual relaxation in (1.10), a problem formulation involving the continuous sharing of subcarriers (in time or frequency) among various bit and power-allocations (e.g., among users) was suggested for orthogonal frequency-division multiple access (OFDMA) systems in [73, 79, 80, 140, 158, 176, 190, 200] and for interference-limited systems in [56, 187]. For time-sharing[3] we consider all allocations $\mathbf{p}^{c,i} \in \mathcal{Q}^c$ indexed by $i \in \mathcal{I}^c = \{1, \ldots, |\mathcal{Q}^c|\}$ and the fractions of time $0 \leq \xi_i^c \leq 1$ that allocation i is used on subcarrier $c \in \mathcal{C}$. Considering time-average objective and constraint values we write the continuous time-sharing relaxation of (1.8) as the linear program (LP)

$$P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,\text{ts}} = \underset{\xi_i^c \geq 0, i \in \mathcal{I}^c, c \in \mathcal{C}}{\text{minimize}} \quad \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}^c} f^c(\mathbf{p}^{c,i}, \hat{\mathbf{w}}, \breve{\mathbf{w}}) \xi_i^c \tag{2.3a}$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}^c} \mathbf{r}^c(\mathbf{p}^{c,i}) \xi_i^c \ \succeq \ \mathbf{R}, \tag{2.3b}$$

$$\sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}^c} \mathbf{p}^{c,i} \xi_i^c \ \preceq \ \hat{\mathbf{P}}, \tag{2.3c}$$

$$\sum_{i \in \mathcal{I}^c} \xi_i^c \ = \ 1, \quad \forall c \in \mathcal{C}, \tag{2.3d}$$

where the constraint in (2.3d) ensures that the time-shares on each subcarrier sum to one. We emphasize that the problem in (2.3) allows for inter-user interference using a finite set of rates according to (1.3), but under a varying fraction of time the possible power combinations on each subcarrier are applied. While we employ time-sharing on a per-subcarrier basis, note that it can also be performed in "aggregated" form [132] over the $\prod_{c \in \mathcal{C}} |\mathcal{Q}^c|$ feasible sum-power and corresponding sum-rate allocations. We have the following result characterizing the optimum of (2.3):

Theorem 1. *Define* $\mathcal{I}_{\boldsymbol{\xi}}^c = \{i \in \mathcal{I}^c | \xi_i^c > 0\}$, *for* $c \in \mathcal{C}$, *and the set of subcarriers where time-sharing occurs as* $\mathcal{C}_{\boldsymbol{\xi}}^+ = \{c \in \mathcal{C} | \ |\mathcal{I}_{\boldsymbol{\xi}}^c| \geq 2\}$. *Assuming feasibility of (1.10) it holds that* $P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,ts} = D_{(\mathbf{R},\hat{\mathbf{P}})}^*$ *and there exists a solution* $\tilde{\boldsymbol{\xi}}$ *to (2.3) with* $|\mathcal{C}_{\tilde{\boldsymbol{\xi}}}^+| \leq 2U$, *i.e., time-sharing is required on at most* $2U$ *subcarriers.*

See Appendix A.1 for a proof.

From Theorem 1 follows also a standard interpretation of Lagrange relaxation in integer programming theory [124] as an optimization over the convex hull of all (discrete) solutions which satisfy the constraints that were *not relaxed*. A very similar intuition of Lagrange relaxation will be given in the following Section 2.1.3, cf. [12, Ch. 5], where the convex hull is taken over a set which includes the objective as one dimension.

---

[3]Note that ultimately we regard time-sharing solely as an algorithmic detour to obtain solutions for our original problem in (1.8), as will become clearer in Section 4.1.

In [187] a scheduling strategy termed "periodic scheduling" was investigated which allows to adapt the target-rates in a multi-carrier system during a scheduling frame of N DMT symbols while meeting an average target-rate per symbol, with the goal to further reduce interference among users. This strategy was shown to be equivalent to a DSM problem over the scheduling frame. Since allocations can only vary from symbol to symbol, periodic scheduling is a special case of time-sharing where the weights $\boldsymbol{\xi}$ are supposed to be taken from a discrete set $\{0, \frac{1}{N}, \ldots, 1\}$. Hence, the continuous time-sharing relaxation in (2.3) is a relaxation of periodic scheduling in this sense. However, by Theorem 1 the dual problem in (1.10) also lower-bounds the optimal, average periodic scheduling objective, meaning that the average gain in terms of objective value by periodic scheduling is bounded by the duality-gap ζ. We refer to [187] for an alternative proof thereof and to [183] where a decomposition algorithm was proposed for the joint periodic scheduling and DSM problem.

A special case is the single-user bit-allocation and power-allocation problem which due to the discreteness of the optimization variables remains a non-convex problem. The pure sum-power minimization problem however is known to be solvable by a greedy bit-loading algorithm [27]. The following theorem shows that no improvement in optimal objective can be obtained by time-sharing in the single-user case.

Theorem 2. *For $U = 1$ it holds that $P^*_{(\mathbf{R}, \hat{\mathbf{P}})} = P^{*, ts}_{(\mathbf{R}, \hat{\mathbf{P}})} = D^*_{(\mathbf{R}, \hat{\mathbf{P}})}, \ \forall \mathbf{R} \in \{\tilde{\mathbf{R}} \in \mathbb{R}^U | \tilde{R}_u = k \cdot \theta, k \in \mathcal{Z}_+, \forall u \in \mathcal{U}\}.$*

See Appendix A.2 for a proof.

A consequence from the proof of Theorem 2 is that in the single-user case greedy bit-loading is optimal for the objective in (1.8a), extending corresponding previous results [27] to a more general objective function.

For completeness, we proceed by analyzing the scalability of the time-sharing problem in (2.3). The number of feasible bit and power-allocations $|\mathcal{Q}^c|, c \in \mathcal{C}$, and therefore the number of variables in (2.3) grows with an increasing number of users U. However, the following result indicates how interference among users restricts this growth.

Theorem 3. *Assuming $\frac{H^c_{ui}}{H^c_{uu}} \geq \alpha > 0$, the number of feasible allocations $|\mathcal{Q}^c|$ on subcarrier $c \in \mathcal{C}$ grows at most polynomially as $O(U^{\hat{U}})$, where the constant exponent $\hat{U}$ is given by*

$$\hat{U} = 1 + \left(\Gamma(2^\theta - 1)\alpha\right)^{-1}. \tag{2.4}$$

See Appendix A.3 for a proof.

The parameter α has the interpretation of a minimal normalized cross-channel attenuation coefficient in the network. Using this minimal value we obtain a lower bound for the (normalized) interference noise per interfering user in (1.3). Assuming all users transmit at a positive rate we thereby obtain an upper-bound for the
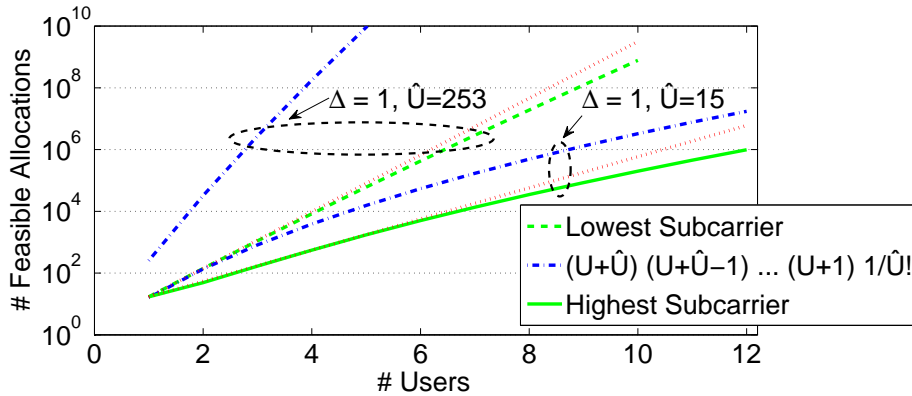
Figure 2.1: Number of feasible bit-allocations $|\mathcal{Q}^c|$ on selected subcarriers $c \in \mathcal{C}$ in symmetric U-user VDSL scenarios with 800 m long lines.

number of users that can be supported by the system in (2.4), hence limiting the growth of the time-sharing formulation in (2.3) as shown by Theorem 3.

Next we will illustrate this bound and the true number of power-allocations $|\mathcal{Q}^c|$ by a DSL example, noting that normalized crosstalk coefficients in DSL networks have been reported to be fairly weak, e.g., $\alpha < -11.3$ dB on typical VDSL lines [28]. In Figure 2.1 we plot the number of possible bit-allocations on the lowest (at about 3 MHz) and highest subcarrier (at about 12 MHz) for a symmetric VDSL upstream scenario with line-lengths $l_u = 800$ m, $\forall u \in \mathcal{U}$, and simulation parameters as specified in Section 4.3. Additionally we show a polynomial of degree $\hat{U}$ shifted by $\hat{U}$ given in (A.9c), as suggested by our bound in Theorem 3. Interference among users clearly decelerates the complexity growth in U. Due to higher crosstalk couplings this effect of interference on complexity is more visible at higher subcarriers where the number of possible allocations is anyway lower due to the more attenuated direct channel. The conclusion we can draw from this example is that while interference has an impact on how fast the number of variables $\sum_{c \in \mathcal{C}} |\mathcal{Q}^c|$ in the LP in (2.3) grows with the number of users U, this number (although of complexity theoretic interest) is too large for a direct LP solution of the time-sharing problem in (2.3). Hence, in Section 4.1 we propose a DSM algorithm based on a decomposition scheme which works with a small subset of these time-sharing variables.

We conclude this section with a result on the complexity of the dual subproblems in (1.12).

Corollary 1 (of Thm. 3). *The per-subcarrier subproblems in (1.12) have polynomial complexity in the number of users U given the assumptions of Theorem 3.*

Proof. By Theorem 3 we have that $|\mathcal{Q}^c|$ has polynomial size in U, and the evaluation of the objective in (1.9) has polynomial complexity. We refer to Chapter 3 for an optimal branch-and-bound scheme for the subproblems in (1.12) with polynomial
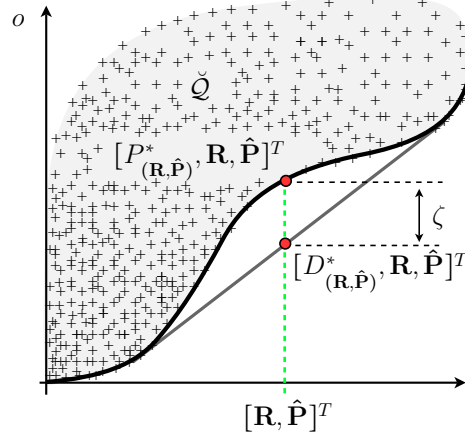
Figure 2.2: Schematic illustration of the "objective-over-constraints" set $\breve{\mathcal{Q}}$ and the duality-gap ζ.

complexity in U, including both, the evaluation of feasible as well as infeasible power-allocations. This concludes the proof. $\qquad\square$

2.1.3 Graphical Interpretation of DSM Problem Relaxations

An intuitive way of rewriting the original problem in (1.8) is by enumerating all possible *sum*-power and *sum*-rate combinations (referred to as "aggregated" form in [132]) and formulating it as the binary optimization problem of selecting the combination with the lowest objective [182]. More precisely, the problem is to pick the optimal element out of the set $\breve{\mathcal{Q}}$, where

$$\breve{\mathcal{Q}} = \sum_{c \in \mathcal{C}} \breve{\mathcal{Q}}^c, \quad \breve{\mathcal{Q}}^c = \left\{ \left[f^c(\mathbf{p}^c, \hat{\mathbf{w}}, \breve{\mathbf{w}}), \mathbf{r}^c(\mathbf{p}^c), \mathbf{p}^c \right]^T | \mathbf{p}^c \in \mathcal{Q}^c \right\}, \qquad (2.5)$$

which we term "objective-over-constraints set" for obvious reasons. A schematic example of this set of points is provided in Figure 2.2. The geometric "min-common/max-crossing" framework derived in [12, Ch. 5] shows that the duality-gap (the difference between $P^*(\mathbf{R}, \hat{\mathbf{P}})$ and $D^*(\mathbf{R}, \hat{\mathbf{P}})$) can be depicted as shown in Figure 2.2, where the dual optimal objective $D^*(\mathbf{R}, \hat{\mathbf{P}})$ is the objective value of the point on the boundary of the polytope $\text{conv}(\breve{\mathcal{Q}})$ with sum-rate/sum-power $(\mathbf{R}, \hat{\mathbf{P}})$. The central observation for the analysis in [12, Ch. 5] is that the dual function in (1.11) can be written as

$$q(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \min_{[o, \breve{\mathbf{r}}, \breve{\mathbf{p}}]^T \in \breve{\mathcal{Q}}} \left\{ [1, -\boldsymbol{\lambda}^T, \boldsymbol{\nu}^T] \cdot \left[\, o, \breve{\mathbf{r}}^T, \breve{\mathbf{p}}^T \, \right]^T + \boldsymbol{\lambda}^T \mathbf{R} - \boldsymbol{\nu}^T \hat{\mathbf{P}} \right\}, \qquad (2.6)$$

which is a linear function in the elements of $\breve{\mathcal{Q}}$. Furthermore, we have that

$$\min_{[o,\breve{\mathbf{r}},\breve{\mathbf{p}}]^T \in \text{conv}(\breve{\mathcal{Q}})} \left\{ [1, -\boldsymbol{\lambda}^T, \boldsymbol{\nu}^T] \cdot \left[o, \breve{\mathbf{r}}^T, \breve{\mathbf{p}}^T \right]^T + \boldsymbol{\lambda}^T \mathbf{R} - \boldsymbol{\nu}^T \hat{\mathbf{P}} \right\} \tag{2.7}$$

$$= \min_{\substack{\boldsymbol{\alpha} \in \mathbb{R}_+^{2U+2}, \sum_{j=1}^{2U+2} \alpha_j = 1, \\ \mathbf{s}^i \in \breve{\mathcal{Q}}, 1 \le i \le 2U+2}} \left\{ \sum_{i=1}^{2U+2} \left(\alpha_i [1, -\boldsymbol{\lambda}^T, \boldsymbol{\nu}^T] \mathbf{s}^i \right) + \boldsymbol{\lambda}^T \mathbf{R} - \boldsymbol{\nu}^T \hat{\mathbf{P}} \right\} \tag{2.8}$$

$$\ge q(\boldsymbol{\lambda}, \boldsymbol{\nu}), \tag{2.9}$$

where in (2.8) we use Carathéodory's theorem [12, p. 696] and in (2.9) we use (2.6), and the reverse inequality holds as $\breve{\mathcal{Q}} \subseteq \text{conv}(\breve{\mathcal{Q}})$. Hence, the minimum value over $\breve{\mathcal{Q}}$ in (2.6) is the same as the minimum value over $\text{conv}(\breve{\mathcal{Q}})$, adding another interpretation of Theorem 1 and Figure 2.2.

2.1.4 Continuous Bit-Loading with Frequency-Division Constraints

While our focus is on the DSM problem for DSL in (1.8) with inter-user interference, it is of additional value to consider a common assumption in wireless networks [80, 114, 158, 176], namely that of assigning subcarriers to users exclusively, i.e., $p_u^c p_i^c = 0, \forall i \in \mathcal{U} \setminus \{u\}, \forall u \in \mathcal{U}$. Denoting the interference-free rates similarly as in (1.3) as $\breve{r}_u^c(p_u^c) = \log_2(1 + \frac{H_{uu}^c p_u^c}{\Gamma N_u^c})$, we write the modified problem of that in (1.8) with *continuous* bit-loading and frequency-division multiple access (FDMA) constraints as

$$\underset{0 \le p_u^c \le \hat{p}_u^c, u \in \mathcal{U}, c \in \mathcal{C}}{\text{minimize}} \sum_{c \in \mathcal{C}, u \in \mathcal{U}} \left(\hat{w}_u p_u^c - \breve{w}_u \breve{r}_u^c(p_u^c) \right) \tag{2.10a}$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} \breve{r}_u^c(p_u^c) \ge R_u, \quad \sum_{c \in \mathcal{C}} p_u^c \le \hat{P}_u, \quad \forall u \in \mathcal{U}, \tag{2.10b}$$

$$p_u^c p_i^c = 0, \quad \forall i \in \mathcal{U} \setminus \{u\}, u \in \mathcal{U}, c \in \mathcal{C}, \tag{2.10c}$$

which is non-convex due to the complementarity (i.e., FDMA) constraints in (2.10c).[4] Conditions for optimality of an FDMA solution in an interference-limited sum-rate maximization problem with continuous bit-loading were derived in [72]. However, these conditions essentially imply a sufficiently strong crosstalk coupling which is not realistic in DSL. The per-subcarrier dual function for this problem,

---

[4]Note that the bit-cap $\hat{\theta}$ maps uniquely to a maximum power constraint through $\breve{r}_u^c(p_u^c)$ and can hence be enforced by an adequate definition of the power mask $\hat{p}_u^c$.

corresponding to (1.12) for our original problem, is, $\forall c \in \mathcal{C}$, given as

$$\breve{q}^c(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \min_{u \in \mathcal{U}} \left\{ \underset{0 \leq p_u^c \leq \hat{p}_u^c}{\text{minimize}} \left((\hat{w}_u + \breve{\nu}_u)p_u^c - (\breve{w}_u + \breve{\lambda}_u)\breve{r}_u^c(p_u^c) \right) \right\}, \tag{2.11}$$

where $\breve{\boldsymbol{\nu}}, \breve{\boldsymbol{\lambda}} \in \mathbb{R}^U$ are the Lagrange multipliers associated with the rate and power constraints in (2.10b), respectively, whereas the dual problem and the dual function in (1.10) and (1.11), respectively, remain unchanged compared to our original problem. The subproblems in (2.11) can be solved efficiently as the objective is concave and (even the problem with discrete rates) can be optimized, for each user separately, by first-order optimality conditions and projection onto the simple power mask bounds in (2.11), cf. Appendix A.7 for a description of this solution approach in a closely related problem. The final task in solving the subproblems in (2.11) is to pick the user with the smallest (optimal) objective value. In total, the complexity of solving the subproblems in (2.11) is $\mathcal{O}(U)$, cf. the corresponding result without FDMA constraints in Corollary 1. As mentioned in Section 2.1.1 the dual problem in (1.10) is concave, so it can be solved in polynomial time using the ellipsoid method, which together with the linear complexity of solving the subproblems in (2.11) means that the dual problem to the FDMA problem in (2.11) has polynomial complexity (both, in the continuous and discrete bit-loading cases), cf. [114] for more details.

As mentioned in Section 2.1.2, another popular relaxation for the FDMA problem is the continuous sharing of subcarriers. Introducing sharing coefficients $\xi_u^c \in \mathbb{R}$, replacing the FDMA constraints in (2.10c), this time-sharing relaxation of the FDMA formulation in (2.10) can be written as

$$\underset{0 \leq p_u^c \leq \hat{p}_u^c, \xi_u^c \geq 0, u \in \mathcal{U}, c \in \mathcal{C}}{\text{minimize}} \sum_{c \in \mathcal{C}, u \in \mathcal{U}} \left(\hat{w}_u \xi_u^c p_u^c - \breve{w}_u \xi_u^c \, \breve{r}_u^c(p_u^c) \right) \tag{2.12a}$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} \xi_u^c \, \breve{r}_u^c(p_u^c) \; \geq \; R_u, \quad \sum_{c \in \mathcal{C}} \xi_u^c p_u^c \; \leq \; \hat{P}_u \quad \forall u \in \mathcal{U}, \tag{2.12b}$$

$$\sum_{u \in \mathcal{U}} \xi_u^c = 1, \quad \forall c \in \mathcal{C}, \tag{2.12c}$$

where we work with average power and rate values, with the exception of the mask constraints in (2.12a) which are enforced at all times. The time-sharing relaxation in (2.12) is a convex problem, as can be seen after the transformation $p_u^c = \tilde{p}_u^c(\xi_u^c)^{-1}$ (related to the sharing of subcarriers in frequency) by the following arguments. The function $\breve{r}_u^c(\tilde{p}_u^c)$ is concave and therefore also the function $\xi_u^c \, \breve{r}_u^c \left(\frac{\tilde{p}_u^c}{\xi_u^c} \right)$ as it is the perspective function of $\breve{r}_u^c(\tilde{p}_u^c)$ [22, p. 39], implying that the relaxation in (2.12) is a (polynomial time solvable [112]) convex problem. Having the results of the previous section in mind it comes now as no surprise that the dual relaxation of the FDMA problem in (2.10) derived above is the Lagrange-dual problem to the time-sharing relaxation in (2.12), where due to convexity of (2.12) strong duality

holds under constraint qualification [22]. The proof follows in analogy to that given for Theorem 1 in Appendix A.1.

2.2 The duality-gap in DSM

So far we have analyzed Lagrange relaxation for DSM and showed the equivalence in terms of objective value to a time-sharing relaxation. Hence, both relaxations incur a duality-gap to the primal optimal objective in (1.8a) due to the non-convexity of the primal problem. Therefore we devote this section to investigate the duality-gap ζ more closely. Our focus is on two important special cases, namely a sum-power constrained sum-rate maximization problem and a sum-rate and sum-power constrained sum-power minimization problem.

2.2.1 Illustrative Examples

The constraint sets of *both*, the primal problem and the dual subproblems in (1.8) and (1.12), respectively, involve integer bit-loading constraints. These integer constraints naturally imply a non-zero duality-gap for non-integer target-rates. Therefore we focus in the following on integer target-rates and bit-loading solutions only. In [146] curves of optimal objective values along a *line* of target-rates were investigated. The non-convexity of such a curve implies a strictly positive duality-gap for certain target-rates. The reverse conclusion does however not hold, cf. the geometric interpretation of dual optimization in Section 2.1.3. In the following we will therefore consider the dependency of the duality-gap on the target-rate.

Figure 2.3(a) depicts the duality-gap ζ for any primal feasible target-rates $\mathbf{R}$ in a scenario with symmetric interference and $C = 2$ subcarriers.[5] Noteworthy, for feasible target-rates the duality-gap in this example can be as much as 93% of the corresponding optimal primal objective. Also, higher gap values seem to be located near to the boundary of the rate-region. An exception occurs at approximately equal target-rates for both users where orthogonal allocations are both, primal and dual optimal.

For weighted sum-rate maximization and sum-power minimization problems with FDMA constraints as in (2.10) these duality-gaps may appear when, changing $\mathbf{R}$ or $\hat{\mathbf{P}}$, the subcarrier assignment to users changes at a primal optimal solution [146]. Differently, in DSL systems we empirically observed that duality-gaps may occur around target-rates where users start to strictly load their bits on different subcarriers, while at rates below these points at least one subcarrier is still shared by both users.

---

[5]The parameters were $\hat{P}_u = \infty, \Gamma = 19, H_c^{uu} = 10^{-1}, N_c^u = 4.3125\,10^{-11}\,\mathrm{mW}, H_c^{ui} = 10^{-4}, i \neq u, \hat{p}_c^u = 10^{-2}\,\mathrm{mW}, \forall c \in \mathcal{C}, \forall u \in \mathcal{U}, \hat{\mathbf{w}} = [0.8, 0.2]^T, \breve{\mathbf{w}} = \mathbf{0}$.
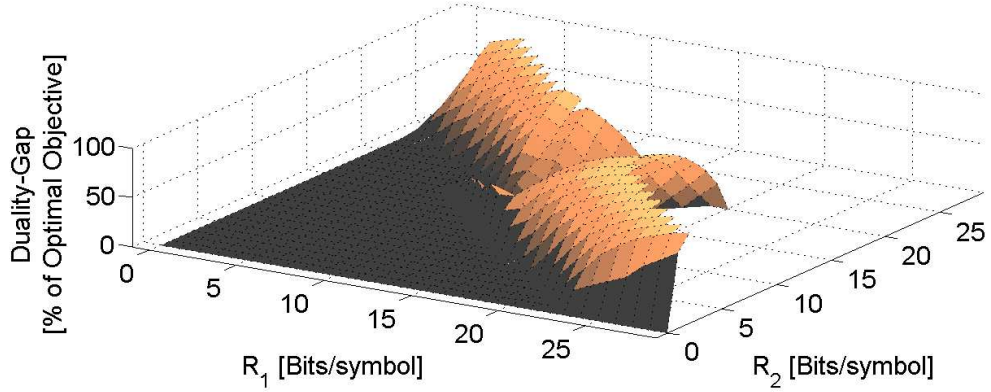
(a)
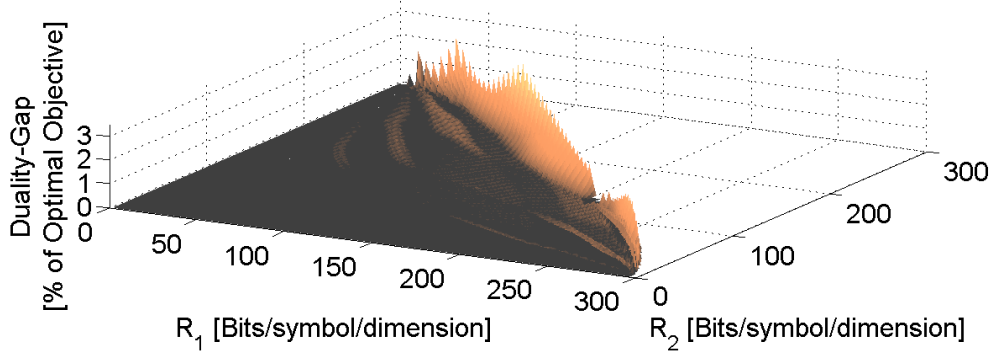
(b)

Figure 2.3: Dependency of the duality-gap on the target-rates $\mathbf{R}$ for the problem in (1.8), a) in a scenario with symmetric interference and $C = 2$ subcarriers, b) in a scenario with $C = 64$ subcarriers, additional sum-power constraints in (1.8c) and orthogonality constraints in (2.10c), and using one-dimensional signal constellations only.

Similar intuitions can be drawn from Figure 2.3(b) where we depict the duality-gap in a system with FDMA constraints and $C = 64$ subcarriers.[6] The maximum duality-gap in this example amounts to approximately $3.2\,\%$ of the corresponding optimal objective value, while it is negligible in most parts of the feasible rate-region.

---

[6]The parameters were $\hat{P}_u = 20\,\mathrm{mW}$, $\mathcal{B}_c^u = \{0, 1, \ldots, \infty\}, \hat{p}_c^u = 1\,\mathrm{mW}$, $\forall u \in \mathcal{U}, \forall c \in \mathcal{C}, H_c^{11}/\Gamma = c^{1.5}, H_c^{22}/\Gamma = (C + 1 - c)^{1.5}, N_c^1 = N_c^2 = 0.1\,\mathrm{mW}$, $\forall c \in \mathcal{C}, \hat{\mathbf{w}} = [1/3, 2/3]^T, \check{\mathbf{w}} = \mathbf{0}$, and only one-dimensional signal constellations were used for simulation complexity reasons, cf. also the similar simulation setup in [146].

2.2.2 Bounds on the Duality-Gap

In [71] it has been shown that the duality-gap in nonconvex problems is zero if the optimal primal solution as a function of the Lagrange multipliers is continuous at all dual optima. This condition implies the convexity of the optimal weighted sum-power objective value in the target-rates. Due to this convexity property it has been argued in [198] that the duality-gap in multi-carrier systems vanishes as the number of subcarriers approaches infinity, cf. also [113] and similar results in the optimization literature [11, Sec. 5.6.1], [4]. Real DSL systems however use at most a few thousand subcarriers, and it is up to now not clear how large the gap is in this case, cf. the simple examples above and also the simulation results in [146] on the duality-gap in *orthogonal* frequency division multiplexing systems with *continuous* bit-loading.

We will develop bounds on the duality-gap in our DSM problem, including one that follows directly from the simple weak-duality relations in (1.13) and (2.1a). This simulation based approach was for example also applied in [114] to an FDMA system. Hence, we emphasize that the key novel aspect of this section lies in analyzing the target-rate dependence of the duality-gap, focussing mainly on interference-limited systems.

In [11, p. 371] the duality-gap of a general non-convex optimization problem was bounded under several assumptions. Based thereon the following can be said about the duality-gap in a (weighted) sum-rate maximization problem with no target-rate constraints.

Corollary 2 (of [11, Prop. 5.26]). *The duality-gap ζ between the optimal objectives of the problem in (1.8) and its partial dual in (1.10) for the special case of a weighted sum-rate maximization problem (i.e., $R_u = \hat{w}_u = 0$, $\breve{w}_u \geq 0, \forall u \in \mathcal{U}$) can be upper-bounded by*

$$\zeta \leq \sum_{u \in \mathcal{U}} \breve{w}_u \cdot \Delta R_u, \tag{2.13}$$

where

$$\Delta R_u = \max_{\{\tilde{\mathcal{C}} \ | \ \tilde{\mathcal{C}} \subset \mathcal{C}, |\tilde{\mathcal{C}}| = (U+1)\}} \left\{ \sum_{c \in \tilde{\mathcal{C}}} \max_{\mathbf{p}^c \in \mathcal{Q}^c} \{r_u^c(\mathbf{p}^c)\} \right\}. \tag{2.14}$$

The result follows from the proof to [11, Prop. 5.26].

Instead of summing over the $(U+1)$ maximum values over subcarriers, we may also upper-bound the right-hand-side in (2.14) using

$$\Delta R_u \leq (U+1) \max_{c \in \mathcal{C}, \mathbf{p}^c \in \mathcal{Q}^c} \{r_u^c(\mathbf{p}^c)\}, \tag{2.15}$$

as done similarly in [11, Prop. 5.26]. As an application example of the bound in
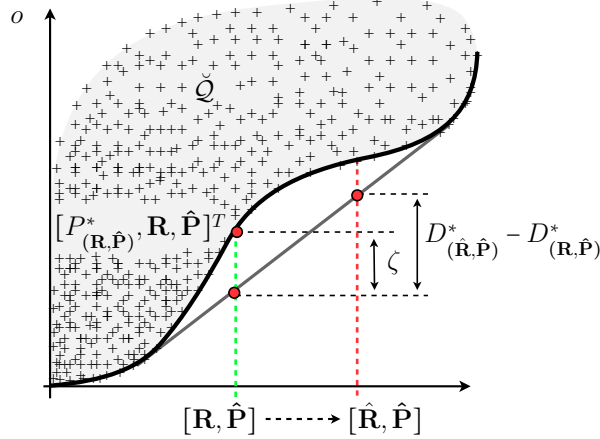
Figure 2.4: Schematic illustration of the discrete set of feasible objective value -
target-rate pairs $\breve{\mathcal{Q}}$, the duality-gap ζ according to the geometric in-
terpretation in [12, Ch. 5], and its bound in (2.16) for $\breve{\mathbf{w}} = \mathbf{0}$.

Corollary 2, using (2.15) and assuming a maximum value of $\hat{\theta} = 15$ bits per sub-
carrier and a DMT symbol-rate of $4 \cdot 10^3$ symb/sec, the duality-gap per user ζ/U is
independently of the channel parameters and other system constraints bounded by
0.66 Mbps and 3.06 Mbps for $U = 10$ and $U = 50$ users, respectively.

However, the assumptions made in the proof of [11, Prop. 5.26] do not hold for the
target-rate constrained problem in (1.8), as integer bit-loadings and an interference
channel are considered. A modified proof allows to show the following bound on the
duality-gap of the problem in (1.8), cf. Figure 2.4 for the idea behind the bound.

Theorem 4. *The duality-gap ζ between the optimal objectives of the problem in
(1.8) and its dual in (1.10) is upper-bounded by*

$$\zeta \leq D^*_{(\hat{\mathbf{R}},\hat{\mathbf{P}})} + \sum_{u \in \mathcal{U}} \breve{w}_u \Delta R_u - D^*_{(\mathbf{R},\hat{\mathbf{P}})}, \tag{2.16}$$

*where $D^*_{(\hat{\mathbf{R}},\hat{\mathbf{P}})}$ is the optimal cost of a perturbed problem to (1.10) with modified
target-rates $\hat{\mathbf{R}} \in \mathbb{R}^U$,*

$$\hat{R}_u = R_u + \begin{cases} \Delta R_u, & \text{if } R_u > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{2.17}$$

See Appendix A.4 for a proof.

The proof is based on an application of the Shapley-Folkman theorem [11, p.
374] to a feasible solution of a perturbed problem to (1.8) with increased target-
rates, and the boundedness of the set of feasible per-subcarrier bit-loadings in (1.7).

Assuming the maximum number of bits loadable per subcarrier we make the most

conservative estimate of the rate-functions' lack of convexity, cf. [4]. Furthermore, the computation of this bound necessitates to solve the dual problem in (1.10) *twice* with different target-rate vectors. Solving the dual problem only *once* we may already apply the weak-duality relation [22, Ch. 5] in (1.13), which is valid for any *primal and dual feasible* solutions. More precisely, a simple bound available during the optimization process is given by

$$\zeta \leq \sum_{c \in \mathcal{C}} f^c(\tilde{\mathbf{p}}^c, \hat{\mathbf{w}}, \breve{\mathbf{w}}) - q(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}), \tag{2.18}$$

where $\tilde{\mathbf{p}}$ is feasible in (1.8) and $\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}$ are feasible in (1.10). Note that this bound is less computationally complex compared to (2.16) and additionally bounds the sub-optimality of *any* primal feasible solution.

Yet another bound is computable for the special case of a (weighted) sum-power minimization problem without optimizing the dual problem in (1.10), but under feasibility assumptions only.

Theorem 5. *Given target-rates* $\mathbf{R}$*, assume that the problem in (1.8) is feasible under the modified target-rates* $\hat{\mathbf{R}}$ *as defined in (2.17). The duality-gap* ζ *between the optimal objectives of the original problem in (1.8) and its Lagrange-dual problem in (1.10) for the special case of a weighted sum-power minimization problem (i.e.,* $\breve{w}_u = 0, \hat{w}_u \geq 0, \forall u \in \mathcal{U}$*) is then upper-bounded by*

$$\zeta \leq \frac{2}{\theta} \cdot \sum_{u \in \mathcal{U}} \Delta R_u \cdot \max_{c \in \mathcal{C}} \{\Delta p_u^c\}, \tag{2.19}$$

and where

$$\Delta p_u^c = \max_{\substack{\{ \{\mathbf{p}^c, \tilde{\mathbf{p}}^c\} \mid \mathbf{p}^c, \tilde{\mathbf{p}}^c \in \mathcal{Q}^c, \\ r_u^c(\tilde{\mathbf{p}}^c) = r_u^c(\mathbf{p}^c) + \theta, \\ r_i^c(\tilde{\mathbf{p}}^c) = r_i^c(\mathbf{p}^c), \forall i \in \mathcal{U} \setminus u\}}} \left\{ \sum_{u \in \mathcal{U}} \hat{w}_u \left(\tilde{p}_u^c - p_u^c\right) \right\}. \tag{2.20}$$

See Appendix A.5 for a proof.

The proof is based on the weak-duality relation in (1.13) and the application of Theorem 4 for modified target-rates. We emphasize that we kept the bound general by avoiding to make any specific assumptions on the optimal power-allocations to (1.12). While loosening the bound, this in fact maximizes the set of target-rates $\mathbf{R}$ for which it is applicable. Furthermore, we expect this bound to decrease with an increasing number of subcarriers for a constant total bandwidth as it depends on the maximum number of bits loadable per subcarrier in (2.14).
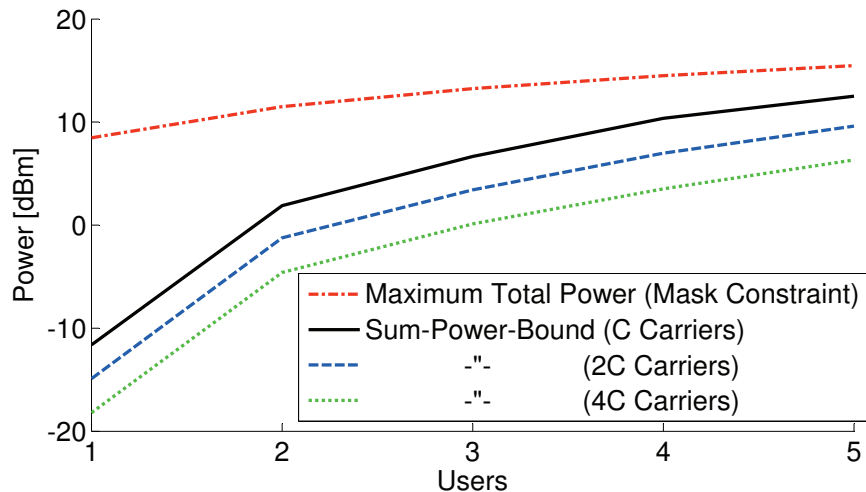
Figure 2.5: Duality-gap bound in (2.19) for different numbers of subcarriers over a constant bandwidth in a DSL scenario with users collocated at 500 m.

Simulations

In this section we will present simulation results for all derived bounds in an upstream very high speed DSL (VDSL) scenario using equal weights $\hat{w}_u = 1$ for all users $u \in \mathcal{U}$. Furthermore, we assume $\breve{w}_u = 0, \hat{w}_u = 1, \forall u \in \mathcal{U}$, i.e., we consider only pure sum-power minimization problems which allows us to directly relate the duality-gap to our transmit-power objective. The simulation parameters were chosen according to the VDSL standard in [50], with an SNR gap $\Gamma = 12.8$ dB, a flat spectral mask constraint at -60 dBm/Hz, two transmission bands as defined in band plan 997-M1x-M, and neglecting the sum-power constraints in (1.8c). The background noise comprised VDSL noise A [50] added to a constant noise floor at -140 dBm/Hz. The first scenario is one where all users $u \in \mathcal{U}$ are collocated at 500 m distance from the deployment point. This case serves solely to investigate the dependencies of the rough bound in (2.19) w.r.t. the number of users and the subcarrier width, cf. Figure 2.5. As can be seen, the bound value increases with the number of users U and eventually exceeds the largest possible value given by summation of the spectral mask constraints over all users and subcarriers. However, as expected at the end of the previous section, by decreasing the subcarrier width (increasing the number of subcarriers for a constant bandwidth) the bound values also become smaller. This is in accordance with the simulation results obtained in [114] for an FDMA system and the corresponding theoretical result showing a decrease of the duality-gap with an increase in the number of subcarriers. This decrease was shown to occur at an asymptotical rate proportional to the inverse of the square-root of the number of subcarriers [114].

Next we consider distributed DSL scenarios with $U = 1$, $U = 2$, or $U = 3$ users
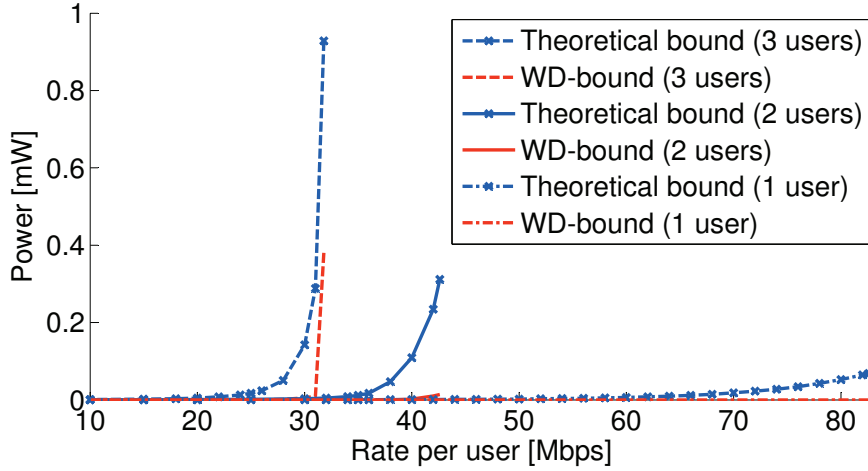
Figure 2.6: Comparison between the duality-gap bound in (2.16) and the weak-duality bound in (2.18) per symbol in a distributed DSL scenario.

located at $\{500\}$ m, $\{400, 600\}$ m, or $\{300, 500, 700\}$ m distance from the deployment point, respectively. Figure 2.6 compares the duality-gap bounds in (2.16) and (2.18) using these 3 scenarios and showing the dependency on the target-rates. We observe that both bounds become looser as target-rates increase, indicating a certain "loss of convexity". Moreover, the theoretical bound in (2.16) turned out to be less tight in this simulation results than the weak-duality bound in (2.18). Inequality (2.19) gives in these scenarios an upper-bound of 0.07, 1.43, and 3.37 mW for $U = 1, \ldots, 3$, respectively. Note that these values are irrespective of the target-rates and roughly up to a factor of 5 higher than the bound in (2.16) from which this bound was derived. Furthermore, the bound in (2.16) gives non-zero values even for $U = 1$, while from Theorem 2 we know that the duality-gap must be zero in this case. Based on the weak-duality bound in (2.18) it is provable for all but the highest feasible target-rates that the duality-gap of the problems studied in Figure 2.6 is vanishingly small. We will see further simulation results which use weak-duality to show a negligible duality-gap in Sections 3.3.5 and 4.3.1.

2.3 Minimizing the Line-Driver Power Consumption

The power consumption of a DSL transceiver can be divided according to its three major parts: the *digital front-end* - performing digital signal processing; the *analog front-end* - performing ADC and DAC as well as signal filtering; and the *line-driver* (LD) - performing power amplification. Depending on the used transmission profile (e.g., bandwidth, maximum aggregate transmit-power) the LD power (LDP) consumption can amount to between $20\,\%$ and $50\,\%$ of the transceiver's total power consumption. The main focus for energy savings in DSL therefore lies on the LDP
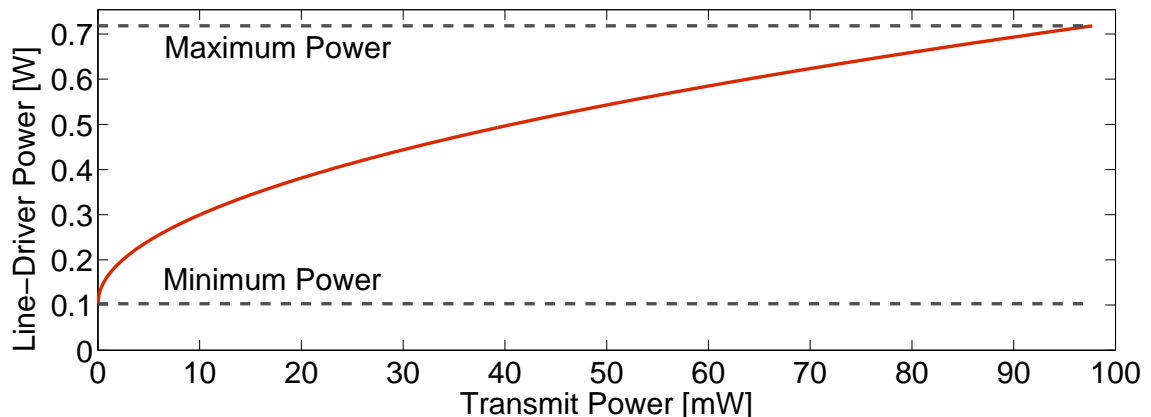
Figure 2.7: ADSL2+ line-driver model from [137] for a maximum aggregate transmit-power of 19.9 dBm.

consumption [66,68].

However, so far we have studied optimization models for DSM targeting a transmit-power (TP) reduction. The TP constitutes roughly between 0.5 % and 10 % of the transceiver's total power budget per port at the network side (depending on technology, profile, and equipment), and even less at the customer side, further depending on the functionality of the home gateway [48]. Therefore an important question is how well the LDP can be optimized by our energy-efficient, hardware-oblivious DSM approach.

2.3.1 Optimization Models

The following investigations are based on the LDP consumption models for different classes of LDs in [66,185], where the LDP as a function of the total TP $P_u = \sum_{c \in \mathcal{C}} p_u^c$ of user u is given in the form of

$$f_u^{LD}(P_u) = \hat{v}_u \sqrt{P_u} + \breve{v}_u, \tag{2.21}$$

and where the parameters $\hat{v}_u \in \mathbb{R}_+$ and $\breve{v}_u \in \mathbb{R}_+$ are dependent on the hardware and system model.[7] The intuition behind the shape of this function is that the sum of the average output power P_u and the average power P_{diss} dissipated in the LD (excluding the quiescent power) can be written as [133]

$$P_u + P_{diss} = \mathbb{E}\{|V_O| \cdot |I_O|\} + \mathbb{E}\{(V_S - |V_O|)\,|I_O|\} = V_S \cdot \mathbb{E}\{|I_O|\} \propto \sqrt{P_u}, \tag{2.22}$$

---

[7]The specific model and parameterization used throughout this thesis is that of a class-AB line-driver specified in the simulation section in [185].

where V_S is the constant supply voltage of the LD, and V_O and I_O are the output voltage and current, respectively. We will use two key features of this function: a) it is monotonously increasing, and b) concave in the sum-power P_u (and hence also in $p_u^c, c \in \mathcal{C}$). The only DSM algorithm for LDP optimization proposed so far [66] uses continuous bit-loading (i.e., $\mathcal{B} = [0, \hat{\theta}]$). Hence, as a comparison to this scheme seems natural when comparing LDP and TP optimization, we will throughout this section focus on the *continuous* bit-loading problem with inter-user interference. Similarly to our original formulation in (1.8) and as in previous studies in the field of DSL [33,130] we mathematically formulate the problem of minimizing the transmit sum-power in DSL in the form

$$J^{\mathrm{TP}} = \underset{p_u^c, u \in \mathcal{U}, c \in \mathcal{C}}{\mathrm{minimize}} \quad \sum_{u \in \mathcal{U}} \sum_{c \in \mathcal{C}} p_u^c \tag{2.23a}$$

$$\text{subject to} \quad \text{Coupling constraints in (1.8b) and (1.8c)} \tag{2.23b}$$

$$0 \le p_u^c \le \hat{p}_u^c, \quad r_u^c(\mathbf{p}^c) \le \hat{\theta}, \quad \forall c \in \mathcal{C}, \forall u \in \mathcal{U}. \tag{2.23c}$$

Similarly, based on the model in (2.21) the problem of minimizing the total LDP consumption in DSL can be stated as

$$J^{\mathrm{LD}} = \underset{p_c^u, u \in \mathcal{U}, c \in \mathcal{C}}{\mathrm{minimize}} \quad \sum_{u \in \mathcal{U}} \sqrt{\sum_{c \in \mathcal{C}} p_u^c} \tag{2.24a}$$

$$\text{subject to} \quad \text{Constraints } (2.23b) \text{ and } (2.23c), \tag{2.24b}$$

where for simplicity of exposition we assume identical LD models for all users which allows us to omit the added constant $\breve{v}_u$ and the factor $\hat{v}_u, u \in \mathcal{U}$, as they have no influence on the optimal solution in this case. Note that the latter factors can easily be reintroduced under the numerical optimization approaches in Section 2.3.3. For instance, heterogeneous LD models will be considered for the simulations in Section 2.3.3.3. For brevity we will denote the optimal per-user sum-power values for the problems in (2.23) and (2.24) by $\mathbf{P}^{\mathrm{TP}} \in \mathbb{R}_+^U$ and $\mathbf{P}^{\mathrm{LD}} \in \mathbb{R}_+^U$, respectively.

2.3.2 Analysis of the Optimization Models

Before turning to the numerical optimization of the problems in (2.23) and (2.24) we analyze their solutions and the difference between their solutions in terms of LDP independently of their exact solution value. To begin with we define the set of possible solutions, referred to as the "power-region".

Definition 1 ("Power-Region"). *The power-region associated with the problems in*

(2.23) and (2.24) is defined as the set

$$\mathcal{P} = \{\mathbf{P} \in \mathbb{R}^U \mid \exists p_u^c, c \in \mathcal{C}, u \in \mathcal{U}, \text{ which are feasible in (2.23)}, P_u = \sum_{c \in \mathcal{C}} p_u^c\}. \quad (2.25)$$

Theorem 6. *The sum-power vectors $\mathbf{P}^{TP}$ and $\mathbf{P}^{LD}$ achieved at a solution of the power minimization problems in (2.23) and (2.24), respectively, both lie on the boundary of the power-region $\mathcal{P}$ as defined in (2.25), i.e., $\nexists \mathbf{P} \in \mathcal{P}, \mathbf{P} \neq \mathbf{P}^{TP}, \mathbf{P} \preceq \mathbf{P}^{TP}$ and $\nexists \mathbf{P} \in \mathcal{P}, \mathbf{P} \neq \mathbf{P}^{LD}, \mathbf{P} \preceq \mathbf{P}^{LD}$.*

Proof. The proof simply follows from the monotonicity of the objectives in (2.23a) and (2.24a), respectively. □

Theorem 6 also suggests a practical heuristic approach for LDP optimization, namely through a sequence of gradient-based updates of sum-powers $\mathbf{P}$ with gradients computed from the objective functions $f_u^{LD}(P_u), u \in \mathcal{U}$, and DSM-based projections onto the power region. We refer to [23] where a similar idea was applied to a rate-utility maximization problem, to Section 4.1.3 where we define a DSM formulation to accomplish the mentioned projections, and to Section 2.3.3.2 for further intuition behind this idea. However, while in [23] a non-concave maximization of a rate-utility function is performed over the rate-region, in (2.24) we face a concave *minimization* problem over the power-region $\mathcal{P}$.

The following theorem identifies the smallest problem instances where a difference between the two problems in (2.23) and (2.24) in terms of LDP may occur, and which we will study further in Section 2.3.3.

Theorem 7. *Differences between the optimal solutions of the problems in (2.23) and (2.24) in terms of LDP can only occur for $U \geq 2$ and $C \geq 2$.*

See Appendix A.6 for a proof.

Next we define the relative gain by LDP optimization in (2.24) compared to TP minimization in (2.23) as

$$\xi = \frac{\hat{v} \sum_{u \in \mathcal{U}} \left(\sqrt{P_u^{\text{TP}}} - \sqrt{P_u^{\text{LD}}} \right)}{\hat{v} \sum_{u \in \mathcal{U}} \sqrt{P_u^{\text{TP}}} + U \cdot \check{v}}. \quad (2.26)$$

In the following we derive a bound on ξ for any number of users U and subcarriers C with powers p_u^c summing to the total power $P_u = \sum_{c \in \mathcal{C}} p_u^c$. More precisely, we have

$$\sqrt{\sum_{u \in \mathcal{U}} P_u^{\text{TP}}} \leq J^{\text{LD}} \leq \sum_{u \in \mathcal{U}} \sqrt{P_u^{\text{TP}}}, \quad (2.27)$$

where the first inequality holds due to the concavity and monotonicity of the model in (2.21), and the optimality of $\sum_{u \in \mathcal{U}} P_u^{\mathrm{TP}}$ in (2.23), and the second inequality holds due to feasibility of a solution to the problem in (2.23) for the problem in (2.24). Using (2.27) in (2.26) we obtain the bound

$$\xi \leq \frac{\hat{v} \left(\sum_{u \in \mathcal{U}} \sqrt{P_u^{\mathrm{TP}}} - \sqrt{\sum_{u \in \mathcal{U}} P_u^{\mathrm{TP}}} \right)}{\hat{v} \sum_{u \in \mathcal{U}} \sqrt{P_u^{\mathrm{TP}}} + U \cdot \breve{v}}, \tag{2.28}$$

which is only dependent on the solution of the problem in (2.23) and shown in Figure 2.8. Expanding our intuition from Theorem 7, we see that this simple bound does not allow for any LDP reduction by direct LDP optimization in (2.24) compared to the TP minimization in (2.23) when all but one user transmit with low power. Using Jensen's inequality [22, p.77] we have $U \sqrt{1/U \sum_{u \in \mathcal{U}} P_u^{\mathrm{TP}}} \geq \sum_{u \in \mathcal{U}} \sqrt{P_u^{\mathrm{TP}}}$, which applied to (2.28) and after simple reformulations leads to a solution-independent upper bound given by

$$\xi \leq \left(1 - \frac{1}{\sqrt{U}} \right) \left(1 + \frac{\breve{v} \sqrt{U}}{\hat{v} \sqrt{\sum_{u \in \mathcal{U}} \hat{P}_u}} \right)^{-1}. \tag{2.29}$$

The gain ξ for $U = 2$ users is for instance bounded by $1 - 1/\sqrt{U} \approx 30\%$. The bounds in (2.28) and (2.29) are identical when $P_u^{\mathrm{TP}} = \hat{P}_u = P, \forall u \in \mathcal{U}$. However, if the solution to the TP minimization problem in (2.23) demands all users to use maximum sum-power, by sum-power optimality in (2.23) the same must hold in the LDP minimization problem in (2.24) and so the difference between the two must vanish.

In this section we have located the solutions of our two optimization problems on the boundary of a power-region and identified potentially insightful problem instances. In the next section we will use this information to study the real gain ξ by directly optimizing the LDP model through numerical methods.

2.3.3 Empirical Numerical Optimization Study

We will use three approaches to obtain insights into the differences between TP and LDP minimization in terms of the LDP consumption founded on the functional model in (2.21): The first one is based on an efficient but possibly suboptimal successive geometric programming (GP) approximation used in order to identify problem parameters under which differences between the optimal solutions of the two optimization problems occur. While it is known that the TP optimization problem can be approached by GP, our contribution is to recognize this fact for the LDP optimization problem. The second approach is based on the globally optimal solution of both problems in (2.23) and (2.24). Global optimality is a necessary property
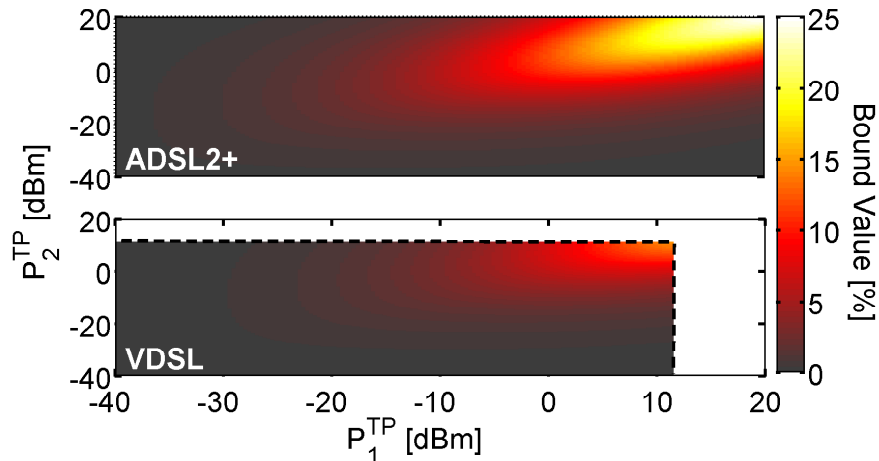
Figure 2.8: Upper-bound in (2.28) on the LDP difference between the solutions of the problems in (2.23) and (2.24) for $U = 2$ users under realistic ADSL2+ and VDSL LDP models with a maximum aggregate TP of 19.9 dBm and 11.5 dBm, respectively.

to study the power-region in Definition 1 and the location of the solutions to the problems in (2.23) and (2.24) in this region. For solving these non-convex and rate-constrained LDP and TP optimization problems we found it necessary to develop a problem-specific algorithm, deviating in various aspects from the approaches proposed for related rate-maximization problems in [49, 192]. The third approach uses the heuristic successive convex approximation algorithms proposed in [68] and [130], respectively. These two algorithms allow to study problem instances of large size and with realistic system parameters.

2.3.3.1 DSM based on successive SINR-approximation and geometric programming

Geometric programs (GP) are a class of problems which is not convex but can easily be converted into a convex form by logarithmic transformations [21]. They were applied to power control in [35, 39], where also successive GP approximations were proposed for non-convex problems based on monomial [21] or SINR approximations [130]. For a short introduction to GP and the corresponding problem transformation of the LDP optimization problem in (2.24) we refer to Appendix C.1.

As mentioned above, our motivation for applying successive GP is to solve numerous small problem instances ($U = C = 2$) in order to identify problem parameters which lead to a substantial gain ξ by LDP optimization compared to TP optimization. We generated numerous problem instances of (2.23) and (2.24) by setting H_{21}^c and H_{12}^c to all combinations out of the set $\{-90, -67.5, -45, -22.5, 0\}$ dB, and for each of these combinations forming all target-rate combinations sampling the users'
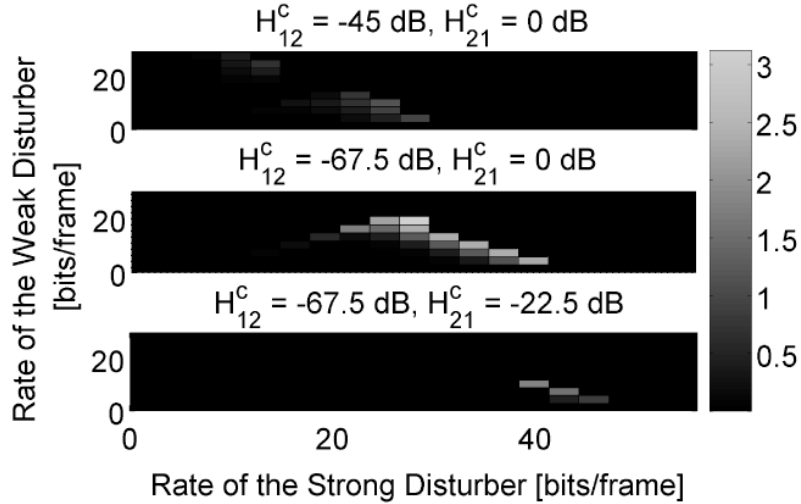
Figure 2.9: Percent gain ξ by direct LDP optimization through the suboptimal successive GP algorithm.

possible rates at 20 equi-distant rate-levels from 0 to the maximum achievable rate (i.e., 400 rate-combinations).[8] After running successive GP for the problems in (2.23) and (2.24) we re-initialize the algorithm with the obtained result for the respective other problem and keep the best solution found for each problem.[9] Also, we multiply the per-user sum-powers by a factor of 500 before applying the LDP model in order to obtain a more realistic estimate of the LDP savings.[10] The result of this experiment can be summarized as follows: Significant values of ξ occurred under unsymmetric settings of target-rates and crosstalk coefficients, especially so when the stronger disturber is the one having the larger target-rate, cf. Figure 2.9. Intuitively this kind of setup results in one user operating with low sum-power (where the derivative of the LDP model in (2.21) is high) while the user with the larger target-rate operates at a higher sum-power (corresponding to a lower derivative of the LDP model in (2.21)). While from a sum-power perspective it may make sense to allow the strong disturber to interfere with the weak disturber due to the specific target-rates, from an LDP perspective the user with the low target-rates is worth protecting more due to the larger derivative of the LDP model at low sum-power values, cf. the LDP model in Figure 2.7. In the next section we select a specific

---

[8]The remaining relevant parameters are $H_{uu}^c = 1$, $\Gamma = 12.3\,\text{dB}$, $\Delta = 4.3125 \cdot 10^3\,[\text{Hz}]$, $N_u^c = 10^{-14} \cdot \Delta\,[\text{mW}]$, $\hat{p}_u^c = 10^{-4} \cdot \Delta\,[\text{mW}]$, $u \in \mathcal{U}$, $c \in \mathcal{C}$, $\hat{\theta} = \infty$.

[9]This sequential re-initialization process is stopped in case the best solution found for both problems does not improve for more than three consecutive iterations.

[10]By multiplication with 500 we heuristically scale the transmit sum-power values to that of a system with 1000 subcarriers in order to obtain LDP values through our LDP model which are somewhat comparable to those under more realistic system parameters in the following sections.

scenario based on these insights for further investigation.

2.3.3.2 Global solutions of non-convex LDP optimization problems using difference-of-convex-functions programming (DCP)

Difference-of-convex-functions programming (DCP) [77] is a widely applicable approach in global optimization, where non-convex objective and constraint functions are reformulated as the difference of convex functions, cf. [49, 192] for recent applications in power control. Similarly to the reformulation shown in [49, 192] for a rate-maximization problem, the rate constraints in (1.8b) can be equivalently written as

$$-\sum_{c\in\mathcal{C}} r_u^c\left(\mathbf{p}^c\right) + R_u = g_u\left(\mathbf{p}\right) - h_u\left(\mathbf{p}\right) \;\leq\; 0, \;\; \forall u \in \mathcal{U}, \tag{2.30}$$

where

$$g_u\left(\mathbf{p}\right) = -\sum_{c\in\mathcal{C}}\log\Big(H_{uu}^c p_u^c + \sum_{j\in\mathcal{U}\backslash\{u\}}\Gamma H_{uj}^c p_j^c + \Gamma N_u^c\Big) + R_u, \tag{2.31}$$

$$h_u\left(\mathbf{p}\right) = -\sum_{c\in\mathcal{C}}\log\Big(\sum_{j\in\mathcal{U}\backslash\{u\}}\Gamma H_{uj}^c p_j^c + \Gamma N_u^c\Big), \; u \in \mathcal{U}, \tag{2.32}$$

are convex fuctions. Writing the objective in (2.24a) formally as $0 - h_0(\mathbf{p})$ with convex function $h_0(\mathbf{p}) = -\sum_{u\in\mathcal{U}}\sqrt{\sum_{c\in\mathcal{C}} p_u^c}$, we can write the problem in (2.24) as the following DCP [77]

$$\underset{p_u^c, u\in\mathcal{U}, c\in\mathcal{C}}{\text{minimize}} \quad -h_0(\mathbf{p}) \tag{2.33a}$$

$$\text{subject to} \quad g_u\left(\mathbf{p}\right) - h_u\left(\mathbf{p}\right) \leq 0, \;\; \forall u \in \mathcal{U} \tag{2.33b}$$

$$\text{Constraints (1.8}c\text{) and (2.23}c\text{).} \tag{2.33c}$$

While in previous applications of DCP in the area of power control [49,192] the problem was in fact solved as a concave minimization problem over a convex constraint set, we have additionally complicating DCP constraints in (2.33b). Correspondingly we developed a more general solution approach, namely a box-based branch-and-reduce algorithm initialized by a successive GP [39] solution, cf. Appendix C.2 for details. Note that this DCP algorithm can similarly be applied to (optimally) solve the TP problem in (2.23).

We use the developed global optimal algorithm to investigate the power region as given in Definition 1. For reasons of tractability we restrict ourselves to a specific scenario ($U = C = 2$) identified using the heuristic in Section 2.3.3.1.[11] In Figure 2.10

---

[11]The relevant selected parameters are that of Section 2.3.3.1 with the exception of $R_1 = 41.36\,[\text{bits/frame}]$, $R_2 = 5.9\,[\text{bits/frame}]$, $H_{12}^c = -67.5\,\text{dB}$, and the value given to H_{21}^c, $c \in \mathcal{C}$.
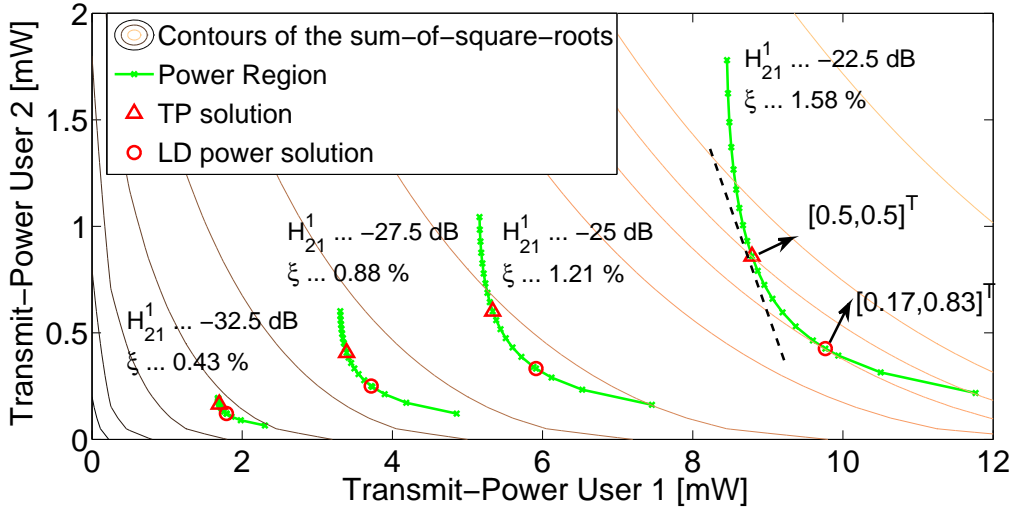
Figure 2.10: Locations of the solutions to the studied problems in (2.23) and (2.24)
in the power region, respectively, and the gain ξ by the latter in terms
of the LDP consumption.

we show the power-regions and the solutions of the problems in (2.23) and (2.24) for
a varying crosstalk parameter H_c^{21}. Firstly we see that both solutions $\mathbf{P}^{\mathrm{TP}}$ and $\mathbf{P}^{\mathrm{LD}}$
lie on the power-region, as predicted by Theorem 6. However, the solutions lie on
different contour lines of the function $\sum_{u \in \mathcal{U}} \sqrt{P_u}$, meaning that they provably differ
in terms of LD power consumption. While the TP solution minimizes $[0.5, 0.5] \cdot \mathbf{P}$
over the power-region, the LD power optimal solution minimizes $[0.17, 0.83] \cdot \mathbf{P}$. In
other words, the LD power optimum is attainable by a weighted sum-power opti-
mization with specific weights. Searching these weights is in fact the idea behind the
projected gradient heuristic indicated in Section 2.3.2. With a decreasing parameter
H_1^{21} the needed sum-powers for constant target-rates decrease, leading to a decrease
of the achievable gain ξ by LD power optimization compared to TP minimization,
cf. Figure 2.10.

2.3.3.3 An experiment in real-sized DSM problems using heuristics

In this section we compare solutions obtained by two DSM heuristics and static
spectrum management (SSM) in terms of their LDP: a) the successive convex ap-
proximation algorithm [186] for the problem in (2.23) which is based on the convex
approximation $\tilde{r}_u^c(\mathbf{p}_c)$ of the rate-function $r_u^c(\mathbf{p}_c)$ as given in Appendix C.1 and in-
troduced in [130] for a rate-maximization problem in DSL; b) the successive LP
approximation algorithm in [68] for the problem in (2.24) which mainly differs from
the above approximation heuristic in that the approximation is linear and the ap-
proximated problems are not solved iteratively but jointly for all users; and c) single-
user water-filling [22] considering a static background noise (i.e., SSM). This static
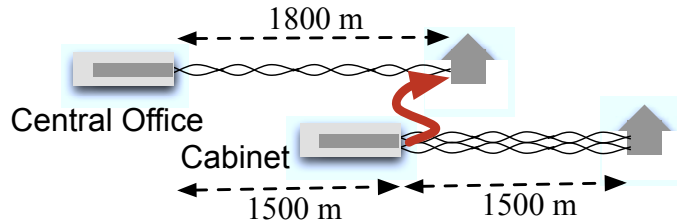
Figure 2.11: Constructed Network Example with 7 cabinet-deployed lines disturbing a CO-deployed line.

background noise includes the highest possible crosstalk noise, assuming that all other systems transmit at the PSD mask level. A novelty we introduce for the comparison of suboptimal DSM algorithms is that after obtaining the result of a DSM scheme we initialize the respective other DSM algorithm with this result and keep the best solutions in terms of LDP and TP objective, respectively.[12] The purpose of this strategy is to avoid the dependency of the comparison on the initialization[13] which might have been chosen in favor of one of the algorithms. The difference to the initialization approach in Section 2.3.3.1 is that we cross-initialize two heuristics, while in Section 2.3.3.1 we applied a single heuristic to two different problems.

Based on the insights of the two previous sections we design a network scenario with realistic parameters where we would expect a difference in LDP between the two considered optimization approaches. This is with respect to the selected channel model (a 99 % worst-case model [59]), the network topology (a near-far scenario with one CO deployed line and 7 cabinet deployed disturbers), the bandplan (showing strong crosstalk with the CO deployed line, see below), the target-rates (low rates for the CO deployed victim line and high rates for the cabinet lines), and the selected DSL systems (the LDP model for the VDSL cabinet lines [185] has a lower slope than that for the ADSL2+ CO line shown in Figure 2.7). More precisely, we consider the near-far downstream scenario shown in Figure 2.11 with 8 lines deployed in the same cable bundle, where 7 VDSL lines are deployed from a cabinet and one ADSL2+ line is deployed from the CO. We set the parameters of the ADSL2+ line in accordance with the standard in [87] (using the non-overlapping bandplan with the integrated services digital network (ISDN) in [87, Annex A]) and of the VDSL lines according to [50] with a total SNR gap of $\Gamma = 12.3\,\mathrm{dB}$ in both systems.[14] The assigned target-

---

[12]This sequential re-initialization process is stopped if no improvement of the best solution found by any of the algorithms was detected for two consecutive iterations.

[13]The PSD for the TP optimization and its first approximation was initialized at a low level of $-120\,\mathrm{dBm}$ per subcarrier and user. The trust-region used in the LDP optimization scheme [68] is set to $-70\,\mathrm{dBm}$ per subcarrier and user, after being initialized with the solution of the sequential TP minimization algorithm [186].

[14]We consider the bandplan setting for fiber-to-the-exchange, mask variant B, and un-notched mask M2, which was chosen due to the high ingress noise into ADSL lines. Thereby we aim to imitate the insightful scenarios found in Section 2.3.3.1.

rates are 1, 2, or 3 Mbps and 10, 13, 16, or 19 Mbps for the CO and cabinet deployed lines, respectively, and we investigate all 12 combinations of these target-rates.[15]

We observed that due to the heuristic nature of both algorithms the LDP optimization did not always give a better total LDP than the TP optimization (corresponding to a negative gain ξ in (2.26)). In summary, the gain ξ in (2.26) was in the studied 12 scenarios between -0.01% and $+0.01\%$. DSM gives a more substantial LDP reduction compared to SSM of between 20% and 40%. While this result is no definite answer to whether or not LDP optimization makes a difference to TP optimization, it is another indication that in practice the difference may be negligible, which motivates the simplification of the optimization in this direction. However, multi-user DSM bares a substantial potential for energy-reduction compared to SSM, as we shall further see in a larger set of scenarios in the following section.

2.3.4 Average Performance Evaluation

Differently to the previous section we will next study the possible LDP reduction by TP optimization (i.e., DSM) compared to SSM in 300 *randomly* generated network topologies with simulation parameters as specified in Section 2.3.3.3. More precisely, we study two deployment scenarios, where the first one consists of 15 ADSL2+ lines with loop-lengths uniformly sampled between 800 m and 1600 m. The second type of scenarios consists of 15 VDSL cabinet-deployed lines with loop-lengths between 300 m and 800 m.[16] We compare the TP optimization algorithm in [186] and the SSM algorithm as described in Section 2.3.3.3. The target-rates are set by multiplying the (scenario dependent) maximum achievable per-user rates as achieved by the heuristic in [68] by factors of $\{0.2, 0.4, 0.6, 0, 8\}$. Differently to above, the crosstalk channel model is based on measured data in [153], where we perform a uniformly random cable selection. Summarizing, the simulation setup does not exaggerate the inter-user crosstalk (e.g., by near-far scenarios or worst-case crosstalk couplings) and therefore provides a realistic evaluation of the energy savings by multi-user DSM compared to SSM.

Next we present the average LDP consumption results together with 99% confidence intervals according to a student t-test. The average LDP consumption in the ADSL2+ scenarios obtained by the sum-rate maximizing DSM algorithm in [68] leads already to an LDP reduction compared to (spectral mask and sum-power constrained) full-power transmission of 38.70% ($\pm 0.97\%$), which has to be compared to the maximum possible savings by TP reduction (which are obtained by reducing the TP to zero) of 85.69%. Hence, even rate-maximizing DSM can be regarded as

---

[15]The maximum rate for the VDSL lines in the considered scenario as found by the LDP optimization algorithm [68] is approximately 19.9 Mbps.

[16]Simulation parameters for both DSL technologies are as specified in Section 2.3.3.3, except that for VDSL we use the bandplan specified in [50] for fiber-to-the-cabinet, mask variant A-M1.
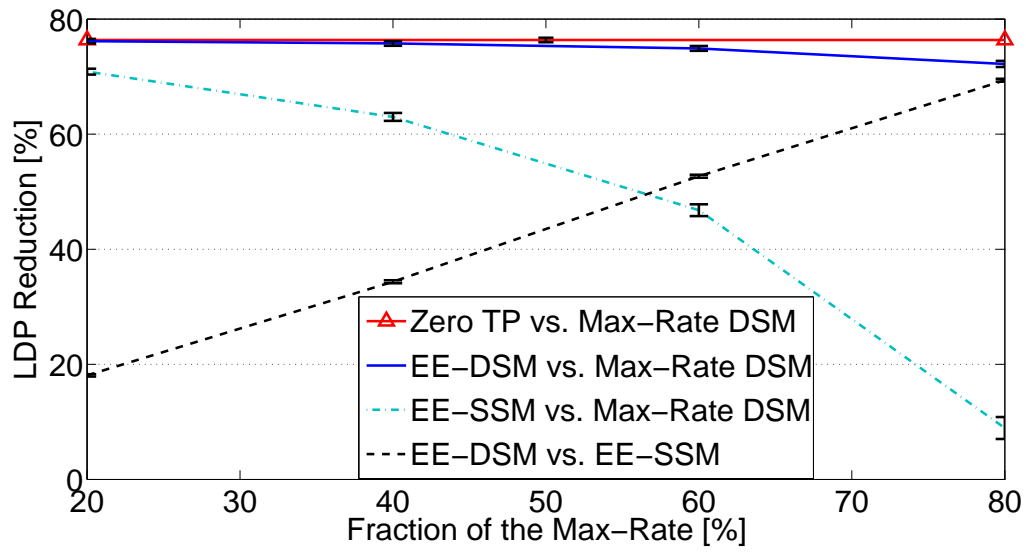
(a)
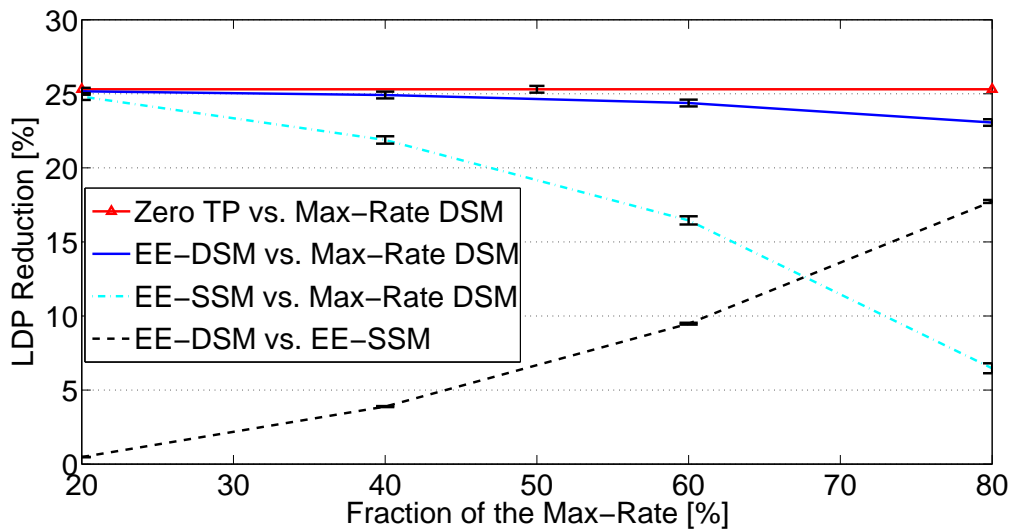
(b)

Figure 2.12: LDP savings achieved by various TP optimization strategies in a) ADSL2+; b) VDSL.

an energy saving technology, as already argued in [41]. In the VDSL scenarios the sum-rate maximization leads to an LDP reduction compared to full-power transmission of 9.10 % (±0.46 %). The maximum possible savings are now only 32.14 %, due

to the lower sum-power constraint as enforced by the spectral mask.

The additional savings by energy-efficient (EE) DSM compared to rate-maximizing DSM are shown in Figure 2.12. In Figure 2.12(a) we see that in the ADSL2+ scenarios multi-user DSM gives (on average) more than 70 % LDP reduction at 80 % of the maximum rates compared to sum-rate maximizing DSM, whereas SSM only results in less than 11 % LDP reduction. Hence, DSM gives substantial improvements compared to SSM, most noticeable at higher rates. In the VDSL scenarios the conclusions are qualitatively similar. However, as shown in Figure 2.12(b), the LDP reduction at 80 % of the maximum rates is now only 23 %, whereas SSM results in less than 7 % LDP reduction.

Estimating the contribution of the LDP in the transceiver's total power budget at between 30 % and 50 % [68, 74], we end up with power savings by energy-efficient DSM in the order of 25 % to 40 % for ADSL2+ and 9 % to 15 % for VDSL compared to full-power transmission, at 80 % of the maximum rates as defined above. This result is in line with the assumed 10 % to 30 % power savings by DSM at the DSLAM in [75].

2.4 Low-Power Modes in DSL

We have so far looked at the case where power could be saved by reducing the set target-rate, and hence the transmit power. Besides the definition of these targets through service level agreements, a second opportunity occurs when modems are enabled to reduce their bit-rates in response to (temporarily) lowered traffic rates.[17] This technique is known as "low-power modes" (LPM) in DSL, and will be briefly surveyed in the following. As this section is mainly meant to provide background information for this chapter and Chapter 7, we refer to [86] for details on the implementation of LPMs, to [189] for an initial study on the optimization of LPMs, to Chapter 7 where we analyze a specific LPM enabling technology by means of various DSM approaches, and to Section 8.3 where we summarize open research directions concerning LPMs.

LPMs are a standardized [86] technique in ADSL2 aiming at reducing the power consumption at the CO side of the DSL link by reducing the downstream transmit rate and power. The initial motivation for LPMs in DSL was the reduction of the heat dissipation and consequently cooling power needed at the CO [42] where typically a large number (up to thousands [75]) of DSL connections are terminated at the DSL access multiplexer (DSLAM). The LPM rate-level was supposed to be just high enough to keep up the basic DSL functionality (e.g., synchronization) and basic telephony services such as the voice-over-Internet protocol (VoIP). Figure 2.13 illustrates the functionality of LPMs in ADSL2 [86]. For example, one may define a

---

[17]We will encounter a third method for defining the target-rates in Chapter 5 where our goal is the maximization of the service coverage.
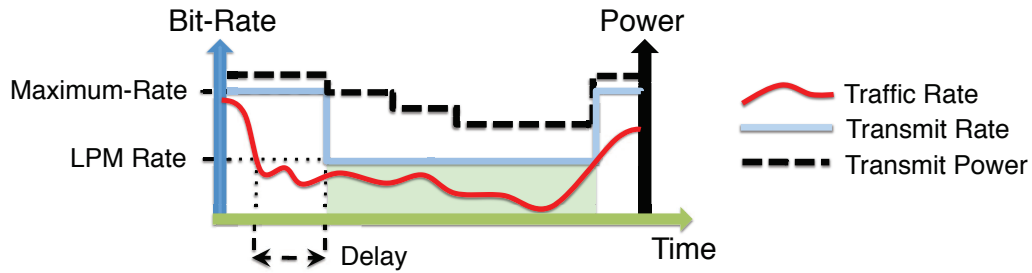
Figure 2.13: Simplified illustration of the LPM functionality in ADSL2 [86].

delay between the time the traffic rate falls below the (single) LPM rate-level and the time the system enters the LPM state. Furthermore, the system exits the LPM instantaneously in order to avoid a user-perceived delay. The exit from the LPM state potentially causes instability in the network as it leads to changes in transmit power and hence in crosstalk noise received on other lines. The effects of LPMs on network stability have been studied in [42], we refer to Chapter 7 for further details on stabilization techniques.

Besides the initial motivation for LPMs, recently energy reduction itself has become an important design criterion in DSL [48,75], having impacts on system scalability (e.g., cooling requirements) and telco's operational expenditure and Carbon dioxide footprint. The European code of conduct on energy consumption of broadband equipment [48] even sets design goals for the energy consumption of DSL equipment. For example, a power reduction by approximately 30 % in the low-power state is foreseen for ADSL2 by 2014.[18] As already stated in Section 1.1, the transceiver's line-driver (LD) accounts for a large part of the transceiver's power budget, e.g., nearly 50 % of the ADSL2 based DSLAM's energy consumption [75]. Furthermore, the LD power (LDP) consumption scales with the transmit power [65,137]. Regarding Figure 2.7 we see that for ADSL2+ transceivers a maximal LDP reduction of 85 % is possible by transmit-power reduction. This conforms to the predicted energy profile for DSLAMs in [20], showing an energy scaling potential of roughly 45 % of the power consumption in full-power state. Differently, in [75] the energy savings by LPMs were more conservatively estimated at 20 %. In [42] maximum savings of 420 mW were found in a specific experimental setup, which assuming a power consumption of an ADSL2 line-card of 1.2 W (the consumption target for 2011-2012 in [48]) corresponds to a saving of 28 %. Furthermore, the saved energy at the transceiver side results in an at least as high extra energy saving in the facility support equipment at the CO, a fact known as the "cascade effect" [127, 141] or "power usage effectiveness" [171]. This makes the scaling of transmit power a viable energy reduction technique. Also, in [141] energy saving modes were found to be the most beneficial energy reduction strategy at the CO, a result based on the energy

---

[18]The precise targets foresee an energy reduction from 3.4 W in the full-power state to 2.4 W on the customer side and from 1.1 W to 0.7 W on the CO side.

reduction targets in [48] and to a large extent in consequence of the cascading effect. This conclusion is further strengthened by the typically low usage of the DSL link. In [20] a network forecast of Telecom Italia for 2015-2020 is given. Therein the link usage (fraction of time the connection is used) and utilization (demanded data-rate compared to the maximum achievable rate) in the wired access network are quantified by 30 % and 10 %, respectively. This indicates the potential for improving the energy-efficiency in DSL by introducing LPMs and adaptive transmission rates. In [51] private DSL and triple-play users are estimated to use the link less than 10 % and 35 %, respectively. Similarly, the study in [144] highlights that the time when the aggregate network traffic in North America is within 5 % of the peak value over the day is only around 2 hours. In [99] the current average Internet usage per day for Bavaria (Germany) was even reported to be as low as 37 minutes (2.6 %).

To obtain an estimate of the saving potential of LPMs we assume a power consumption of an ADSL2 line-card of 1.2 W (the consumption target for 2011-2012 in [48]), an average saving in LDP consumption of 64 %,[19] a share of the LDP in the line-card's power budget of 50 % [69], and a multiplicative energy-saving factor of 2 due to the power usage effectiveness [75, 127, 141]. Altogether we obtain an energy saving potential of 6.7 kWh (or 0.67 Euro assuming an energy price of roughly 10 cent/kWh [136]) per year and DSL line. Summarizing, the efficiency of LPMs depends on the time operated in LPM and the transmit power spent in LPM. In [189] we optimize the LPM rate level for a single DSL line and based on different models and assumptions of the arrival traffic. This bottom-up approach provides a comparable range for the average LDP saving potential compared to that assumed in our estimate (64 %), namely roughly between 60 % and 80 %. An assumption for the work in [189] is that the DSL line is stabilized and the fluctuations in received crosstalk noise masked by the considered increased background noise, cf. Chapter 7 for a specific means of line stabilization.

---

[19]This number is based on an average link usage of 20 % and a potential LDP reduction in LPM during idle-times of 80 % [137], resulting in an average saving potential of $(1-0.2) \times 0.8 = 0.64$.

Part II

Novel DSM Algorithms

Chapter 3 Optimal Discrete-Rate DSM

Our main contribution in this chapter is the proposal of optimal methods for solving the decomposable dual multi-user subproblems in (1.12), and the analysis of the suboptimality of greedy multi-user bit-loading [103] based on these methods. This work does not depend on the specific algorithm which generates the Lagrange multipliers $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$, cf. the Lagrange-dual problem in (1.10).

Combinatorial multi-user rate and power-allocation as in (1.12) is of interest not only in interference-limited DSL systems, but also in other application domains where interference plays a role, such as wireless multi-user networks [36, 105, 135]. A dual optimal multi-user DSM algorithm in DSL with discrete power levels was first proposed in [33], where an exhaustive search was applied to solve the per-subcarrier problems. In [188] we propose a robust extension to DSM where the exhaustive search is made more efficient by avoiding the evaluation of infeasible rate combinations. A branch-and-bound (BnB) algorithm for optimal per-subcarrier search was proposed in [163], where the BnB search had an exponential complexity in terms of time and memory, and ADSL downstream scenarios were simulated for up to 8 users. Optimal *continuous* power-allocation has been studied in [138, 162, 177, 192], where in [192] dual optimal results were presented for up to 8-user VDSL scenarios. In [100] a power control scheme for finite-rate OFDMA systems with a single (base-station) sum-power constraint was proposed. Its peculiarity is the usage of the actual bisection search interval in the dual master problem in order to declare optimality of a per-subcarrier solution (i.e., the subcarrier allocation to a single user and its transmit power) early on in the dual search. Our application of BnB methods to the interference-limited and multi-user per-subcarrier problems also allows to incorporate dual master problem information, as we shall see in Section 3.1 and Chapter 4.

A traditional approach for discrete-rate DSM is discrete bit-loading (DBL) [59]. Various optimal DBL algorithms have been proposed in the single-user case, cf. [5, 27] and the references therein. However, only few heuristic have been described in the multi-user case [103, 184, 196]. To the best of our knowledge no study has up to now investigated the precise suboptimality of multi-user bit-loading schemes in a large set of scenarios.

We propose a low-complexity *optimal* discrete-rate allocation method for the per-subcarrier power control problems in (1.12) which consists of problem-specific im-

plementations of two mechanisms which are generic parts of a modern nonlinear discrete problem solver [8, 143]: a BnB framework and a variable-range reduction technique. We propose in Section 3.1 two BnB schemes which show favorable computational and memory complexity to that previously proposed in [163]. The key feature of our BnB schemes is that maximum bit-loading information is passed between neighboring nodes in the BnB search tree in order to reduce the number of infeasible rate evaluations. In Section 3.2 we suggest a low-complexity *optimal* search-space reduction (SSR) scheme based on a partly interference-free, convex relaxation of the per-subcarrier problems. In the context of related integer programming literature [9, 26, 142, 143] our SSR scheme can be interpreted as a low-complexity, relaxation specific, nonlinear, objective-based [142] variable range reduction technique. SSR is seen to be most effective in scenarios with low levels of crosstalk, e.g., low-bandwidth DSL systems or DSM problems with low target-rates. Simulation results are presented in Section 3.3, where we motivate our problem-specific approach by comparison to a general-purpose solver for mixed-integer non-convex problems [8], investigate the effectiveness of the proposed techniques individually, and demonstrate the effect of target-rates on the solution complexity in a 16-user ADSL2 scenario. Furthermore, we analyze the sum-rate performance of a classical greedy multi-user DBL algorithm [103] for which we provide for the first time precise suboptimality figures in thousand ADSL2 networks using measured channel data from [153]. We will now study the problem in (1.12) in the general multi-user case and drop the subcarrier index c throughout the rest of this chapter for ease of notation. Also, to facilitate an intuitive understanding of the algorithms proposed in the following we rephrase the problem in (1.12) in terms of bit-rates as

$$\underset{\mathbf{r} \in \mathcal{Q}^{\mathbf{r}}}{\text{minimize}} \ f(\mathbf{p}(\mathbf{r}), \hat{\mathbf{w}} + \boldsymbol{\nu}, \breve{\mathbf{w}} + \boldsymbol{\lambda}). \tag{3.1}$$

The complete proposed methods can be found in the tables Algorithm 11 and Algorithm 12. We emphasize that our contribution does not depend on the specific algorithm which generates the Lagrange multipliers $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ and is therefore also applicable to previous work on DSM algorithms. However, for our simulations in Section 3.3 we use the finitely converging spectrum balancing framework described in Chapter 4. It is based on a linear problem (LP) which is iteratively updated using the per-subcarrier solutions of the problems in (3.1). The dual solution of this LP comprises the Lagrange multipliers which are used to define the subproblems in (3.1).
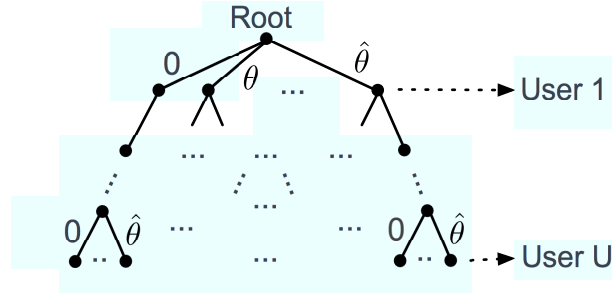
Figure 3.1: Schematic illustration of the search-tree associated with a single per-subcarrier problem in (1.12).

3.1 Branch-and-Bound (BnB) for Discrete-Rate Power Control

Solving the per-subcarrier problem in (3.1) can be regarded as a sequence of U consecutive decisions, each assigning a certain number of bits to one of the U users. This can be illustrated in form of a decision tree, where decisions are made starting with that of user 1, cf. Figure 3.1. A node at tree-depth u corresponds to a vector of bit-allocations $\mathbf{r}^{(u)} \in \mathbb{R}^u$ of *"already loaded users"* 1 up to u. We define the set of *all* users' possible rates (the leaf nodes at the bottom of the search-tree) per subcarrier as $\mathcal{L} = \prod_{u \in \mathcal{U}} \mathcal{B}$, the set of *feasible* rates by $\mathcal{Q}^{\mathbf{r}} = \{\mathbf{r} \in \mathcal{L} | \mathbf{p}(\mathbf{r}) \in \mathcal{Q}\}$, and the set of *infeasible* rates by $\bar{\mathcal{Q}}^{\mathbf{r}} = \{\mathbf{r} \in \mathcal{L} \mid \mathbf{r} \notin \mathcal{Q}^{\mathbf{r}}\}$. Note that this work can be readily extended to more general (finite) sets $\mathcal{B}$. Branch-and-bound (BnB) [101] is a systematic and exact search method for finding the leaf-node $\mathbf{r}^{(U)} \in \mathcal{Q}^{\mathbf{r}} \subseteq \mathcal{L}$ with minimum objective value $f(\mathbf{p}(\mathbf{r}^{(U)}), \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda})$, i.e., for solving the subproblem in (3.1). The algorithm starts at the root node of the tree (cf. Figure 3.1) and at each node makes a branching decision, i.e., it decides which node to explore next at the next-lower tree-level. The second key component of the method beside the branching rule is the computation of lower-bounds on the objective values of any leaf-node belonging to a subtree rooted at node $\mathbf{r}^{(u)}$, cf. Section 3.1.1. However, the *exact* computation of the tightest possible lower bound would necessitate the solution of the problem (cf. (1.3), (1.7), (1.9), and (3.1))

$$\underset{\substack{r_i \in \{0, \theta, \ldots, \hat{\theta}\}, u+1 \leq i \leq U, \\ p_i \in [0, \hat{p}_i], \forall i \in \mathcal{U}}}{\text{minimize}} \sum_{\{i \in \mathcal{U} \mid i \leq u\}} \left((\hat{w}_i + \nu_i)p_i - (\check{w}_i + \lambda_i)r_i^{(u)} \right) +$$

$$\sum_{\{i \in \mathcal{U} \mid i \geq (u+1)\}} ((\hat{w}_i + \nu_i)p_i - (\check{w}_i + \lambda_i)r_i) \tag{3.2a}$$

$$\text{subject to} \quad r_i^{(u)} \leq r_i(\mathbf{p}), \quad 1 \leq i \leq u, \tag{3.2b}$$

$$r_i \leq r_i(\mathbf{p}), \qquad u+1 \leq i \leq U, \tag{3.2c}$$

which differs from the original problem in (3.1) only in the fact that the rates of users $i, i \leq u$, are fixed at $\mathbf{r}^{(u)}$. Consequently, solving (3.1) with a BnB algorithm using the tightest lower bound would be as costly as solving (3.1) by a brute-force enumeration of $\mathcal{Q}^{\mathbf{r}}$. However, the efficiency of the BnB algorithm comes from the use of efficiently computable lower-bounds to the optimal objective of the problem in (3.2), which can be used to infer the suboptimality of a subtree ("pruning" the subtree) based on the comparison of this lower-bound with the objective value of the best ("incumbent") solution $\mathbf{r}^{(U)} \in \mathcal{Q}^{\mathbf{r}}$ found so far. The algorithm only stops when it has either visited or pruned all leaves of the tree. Hence, there is a complexity tradeoff in BnB schemes between the computation of lower-bounds and the exploration of the tree.

3.1.1 Computing lower-bounds in BnB schemes

Before proposing explicit BnB methods we will describe the computation of bounds which are less complex than the solution of the problem in (3.2). We generically denote such a lower-bound for the subtree rooted at $\mathbf{r}^{(u)}$ as

$$lb^{\mathbf{r}^{(u)}}(\mathbf{p}^{\min}, \mathbf{r}^{\max}) = (\hat{\mathbf{w}} + \boldsymbol{\nu})^T \mathbf{p}^{\min} - (\check{\mathbf{w}} + \boldsymbol{\lambda})^T \mathbf{r}^{\max}, \quad (3.3)$$

where $\mathbf{p}^{\min}$ and $\mathbf{r}^{\max}$ are chosen appropriately, cf. [163] for a specific BnB scheme dependent choice. A selection which gives a valid lower-bound in (3.3) to the exact lower bound in (3.2) is $p_i^{\min} = p_i(\mathbf{r}^{(u)}), r_i^{\max} = r_i^{(u)}, 1 \leq i \leq u$, and $p_i^{\min} = 0, r_i^{\max} = \hat{\theta}, (u+1) \leq i \leq U$, where $p_i(\mathbf{r}^{(u)}) = p_i(\mathbf{r}), i \in \mathcal{U}$, as defined in (1.3) with $\mathbf{r} \in \mathbb{R}^U$ being formed by extending $\mathbf{r}^{(u)}$ with $(U - u)$ zeros. This selection results in a valid lower-bound as a) the rates of already loaded users are fixed while those of the remaining users are constrained by $\hat{\theta}$; b) the power needed to support the rates $\mathbf{r}^{(u)}$ in (3.2b) monotonously increases with the crosstalk noise and is therefore the lowest when the power of the remaining users is the lowest possible (i.e., 0); and c) the objective in (3.2a) is monotonously increasing in the powers and decreasing in the rates. In Section 3.1.2.1 and 3.2 we will see two less conservative ways of obtaining bounds by modifying $p_i^{\min}$ and $r_i^{\max}$ for users i, where $(u+1) \leq i \leq U$.

Another means of efficiently lower-bounding $f(\mathbf{p}(\mathbf{r}), \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda})$ for all leaf nodes $\mathbf{r}$ in the subtree rooted at $\mathbf{r}^{(u)}$ also derives from the above key observation that the power needed to support the rate $r_i(\mathbf{p})$ in (3.2b) and (3.2c) without interference among users lower-bounds the corresponding values with interference. A relaxation

of the problem in (3.2) in this respect can be formulated as

$$
\underset{\substack{r_i \in \{0, \theta, \ldots, \hat{\theta}\}, \\ p_i \in [0, \hat{p}_i], \, u+1 \le i \le U}}{\text{minimize}} \quad \sum_{\{i \in \mathcal{U} \,\mid\, i \le u\}} \Big((\hat{w}_i + \nu_i) p_i(\mathbf{r}^{(u)}) - (\breve{w}_i + \lambda_i) r_i^{(u)} \Big) +
$$

$$
\sum_{\{i \in \mathcal{U} \,\mid\, i \ge (u+1)\}} \big((\hat{w}_i + \nu_i) p_i - (\breve{w}_i + \lambda_i) r_i \big) \tag{3.4a}
$$

$$
\text{subject to} \quad r_i \le \log_2\Big(1 + \frac{H_{ii} p_i}{\Gamma \tilde{N}_i} \Big), \; u+1 \le i \le U, \tag{3.4b}
$$

which differs from the exact lower-bounding problem in (3.2) in that the powers p_i for users $i, i \le u$, are fixed at the above mentioned lowest values $p_i(\mathbf{r}^{(u)})$ and the constraints in (3.2b) are neglected. Furthermore, the total received noise considered in the rate functions $r_i(\mathbf{p})$ in (3.2c) is lower-bounded by considering only the interference from already loaded users, which by the same argument as above leads to the lower-bound on the received noise given by $\tilde{N}_i = N_i + \sum_{1 \le j \le u} H_{ij} p_j(\mathbf{r}^{(u)})$, for the remaining users $i, (u+1) \le i \le U$. Solving the problem in (3.4) forms the basis of our search-space reduction algorithm in Section 3.2 and its solution with linear complexity in U is detailed in Appendix A.7.

Yet another relaxation of (3.2) in the form of an LP can be obtained in that we first apply a continuous relaxation of the discrete rate variables, and relax (3.2c) as in (3.4b). This relaxation would truly capture the interference induced to the already loaded users through the constraints in (3.2b). However, we find that the improvements of the pruning process in our BnB schemes are outweighed by its complexity and we will therefore not consider it further.

3.1.2 Depth-First BnB (DFB) and Best-First BnB (BFB) Search

The symmetric BnB scheme in [163], which we refer to as the Regular Splitting based BnB (RSB), branches by expanding all sub-trees in parallel. This leads to exponential worst-case memory requirements, making RSB inapplicable for a larger number of users. The second disadvantage, as the simulations in Section 3.3.2 will reveal, is that testing infeasible allocations $\mathbf{r} \in \bar{\mathcal{Q}}^{\mathbf{r}}$ significantly contributes to the search complexity of this method. The BnB schemes presented in the following remedy exactly these two disadvantages by having linear worst-case memory requirements and introducing an explicit mechanism to exploit the impact of some user's bit-allocation on other users.

3.1.2.1 Depth-first BnB

Depth-first BnB (DFB) explores the search-tree starting from the node $\mathbf{r}^{(U)} = \mathbf{0}$, corresponding to the bottom-left node in Figure 3.1. Bits are iteratively increased,
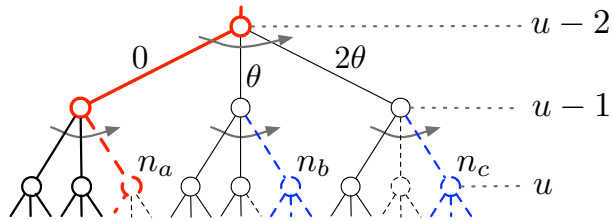
Figure 3.2: Part of a BnB tree under depth-first search for the case of $\hat{\theta} = 2\theta$ containing a set of three neighboring subtrees rooted at level $u-1$; Dashed lines and circles illustrate infeasible allocations; The red path corresponds to the current allocation, based on which we infer infeasibility of the subtrees highlighted in blue.

starting with user U, corresponding to the bottom level of the search-tree. Algorithm 11 summarizes the DFB scheme in more detail, with specific aspects being explained in the remainder of this section. Lines 8–14 and 19–24 implement the searching strategy (i.e., the update of the currently investigated node in the search tree), with an update of the minimal power-allocation $\mathbf{p}^{\mathrm{min}}$ in Line 11 as used in (3.3), and a conditional update of the incumbent objective f^* in Line 13. In Line 19 the sub-tree is pruned based on the comparison of the current lower bound lb to the current upper bounds. In Lines 15–18 we perform further tasks needed to improve the lower bound in (3.3), including an update of $\mathbf{r}^{\mathrm{max}}$ and a call of Algorithm 12 to be specified in Section 3.2.

Information regarding the maximum feasible[1] bit-loading of user u, in a subtree rooted at level $u-1$, can be used to produce an upper bound on user u's feasible bit-loadings in all neighboring subtrees rooted at level $u-1$. This is made possible by the intuitive observation that increasing the rate of any other user $i \in \mathcal{U}$, $i \neq u$, cannot increase the maximum feasible number of bits of user u as increasing the rate of another user can only increase the overall interference. The basic idea is exemplified in Figure 3.2, where the infeasibility of node n_a implies that also the nodes n_b and n_c are infeasible. Consequently, having visited node n_a in the DFB search over the tree illustrated in Figure 3.2 and found this node infeasible, the upper bound on the maximum rate of user u can be set to $r_u^{\mathrm{max}} = \theta$ when searching the neighboring subtrees rooted at level $u-1$. In the general case, having found a node $\mathbf{r}^{(u)}$ infeasible, the DFB can set $r_u^{\mathrm{max}} = r_u^{(u)} - \theta$ in the search of the remaining neighboring subtrees at level $u-1$, i.e., in any remaining subtree rooted at $\mathbf{r}^{(u-1)}$, where $r_i^{(u-1)} = r_i^{(u)}$ for $i = 1, \ldots, u-2$.

Testing feasibility of a node $\mathbf{r}^{(u)}$ requires in general the computation of the unique power-allocation $\mathbf{p}(\mathbf{r}^{(u)})$ as the solution of a set of linear equations (cf. Section 1.2.3).

---

[1]We call a node $\mathbf{r}^{(u)}$ in the search-tree feasible if the extension of the bit-allocation with zero bits for users i, $u < i \leq U$, yields a feasible leaf-node $\mathbf{r}^{(U)} \in \mathcal{Q}^{\mathbf{r}}$. If $\mathbf{r}^{(U)}$ is infeasible (i.e., $\mathbf{r}^{(U)} \notin \mathcal{Q}^{\mathbf{r}}$), no feasible extensions exist.

The value of the observation described above is that by using the upper bound on the rate of user u we can avoid testing certainly infeasible nodes, thus reducing the number of computations required in the BnB implementation. The procedure can be applied at any level of the tree, although the upper bound $r_u^{\max}$ must be re-set to $\hat{\theta}$ when the search back-tracks to level $u - 2$.

Besides lowering the number of feasibility tests performed (cf. the use of variable $\hat{r}_u^{\text{tmp}}$ in Algorithm 11), the reduction of $r_u^{\max}$ also strengthens the lower bound in (3.3), cf. Lines 14 and 18 in Algorithm 11. Furthermore, the initialization of the incumbent in lines 3–5 can be further improved, e.g., by using heuristics [181] or the solutions of other DSM algorithms, cf. Section 3.3. Note that DFB as well as the scheme presented in the following section can make use of an available objective bound Φ to improve the pruning process, cf. Line 19 in Algorithm 11 and [179] where the specific relaxation gives such a bound as a byproduct. The variable f^{lb} in Algorithm 11 serves as a certificate in case no feasible solution with objective lower than Φ is found, i.e., as a lower-bound for the objective of any (e.g., pruned) feasible allocation. In the following we analyze the complexity of DFB.

Corollary 3 (of [179, Thm. 2]). *The complexity of the search scheme DFB in Algorithm 11 for solving the subproblem in (3.1) is polynomial in the number of users U given $\frac{H_{ui}}{H_{uu}} \geq \alpha > 0, \forall u \in \mathcal{U}$.*

Proof. While by Theorem 3 we have that $|\mathcal{Q}^{\mathbf{r}}|$ grows polynomially in U, we also see that the same holds for the complexity per feasible allocation in the BnB Algorithm 11 (i.e., including the solution of a linear system for evaluating $\mathbf{p}(\mathbf{r})$ and the calculation of lower-bounds through (3.3) or the solution of the relaxation in (3.4) as detailed in Appendix A.7). Furthermore, the number of tested *infeasible* allocations $\mathbf{r} \in \bar{\mathcal{Q}}^{\mathbf{r}}$ can be bounded as follows: Assuming a feasible bit-allocation (e.g., $\mathbf{r} = \mathbf{0}$), the algorithm proceeds by increasing the number of bit-steps for user U by one. If this is feasible we obtain another feasible bit-allocation, while if it is not we return to the previous user $U - 1$ and increase its bit-allocation by one, setting that of user U to 0. Following this procedure we find that the maximum number of failed trials per feasible bit-allocation is bounded by the depth of the search-tree U, concluding our argument. □

3.1.2.2 Best-first BnB

Best-first BnB (BFB) differs from DFB in that branching is not performed systematically starting from lower bit-allocations and proceeding to higher ones. More precisely, the algorithm branches the node which has the lowest lower-bound in (3.3). Whether this is a good decision will clearly depend on the quality of the lower-bound. Similarly as in DFB we can make use of the maximum bit-loading information $r_u^{\max}$. However, considering only subtrees rooted at level $u - 1$ with equal bit-allocation $r_i^{(u-1)}$ for all previous levels $i, i \leq u - 2$ (cf. Figure 3.2), in BFB we might visit a
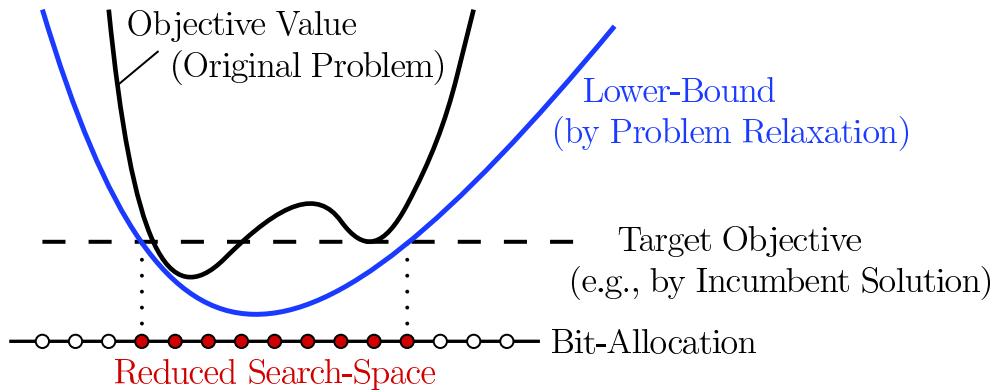
Figure 3.3: Illustration of the idea behind objective-based search-space reduction (SSR).

subtree with higher rate $r_{u-1}^{(u-1)}$ before visiting a subtree with lower rate $r_{u-1}^{(u-1)}$ at level $u-1$. Therefore we might not have access to the closest estimate $r_u^{\max}$ we would have obtained if we had visited the neighboring subtree with lower rate $r_{u-1}^{(u-1)}$ first. This deteriorates the effectiveness of the bounds in (3.3) and the search as a whole. For brevity we omit a detailed algorithm description of BFB as we found BFB inferior compared to DFB for a larger number of users, cf. our results in Section 3.3.

3.2 A Search-Space Reduction (SSR) Scheme

In Theorem 3 an upper-bound on the size of the feasible search space $|\mathcal{Q}^{\mathbf{r}}|$ for the per-subcarrier problem in (3.1) is given which grows with decreasing crosstalk strength. The bound has therefore its highest value if we assume no crosstalk between users. This is however counter-intuitive in face of the fact that a greedy bit-loading algorithm would then solve (3.1) optimally, cf. the proof of Theorem 2 and the fact that the per-subcarrier problem decouples into single-user problems in the absence of interference as seen in Section 3.1.1. In the following we suggest a method which can take advantage of low-interference cases in order to reduce the size of the search-space for the exact problem in (3.2) corresponding to a subtree rooted at $\mathbf{r}^{(u)}$.

The idea behind our SSR scheme is to first solve the lower-bounding problem in (3.4) which partly neglects inter-user interference, and is therefore a relaxation of the exact problem in (3.2). Then, starting from the optimum of the lower-bounding problem, we search the bit-allocations with a lower-bound (i.e., objective value in the lower-bounding problem) below the real objective (i.e., with fully considered inter-user interference) of a known feasible solution for the exact problem in (3.2). This idea is illustrated in Figure 3.3 where it is shown how a lower bound on the objective provides a tightened superset of all bit-allocations that can improve upon the target objective. Correspondingly, the SSR scheme in Algorithm 12 contains two
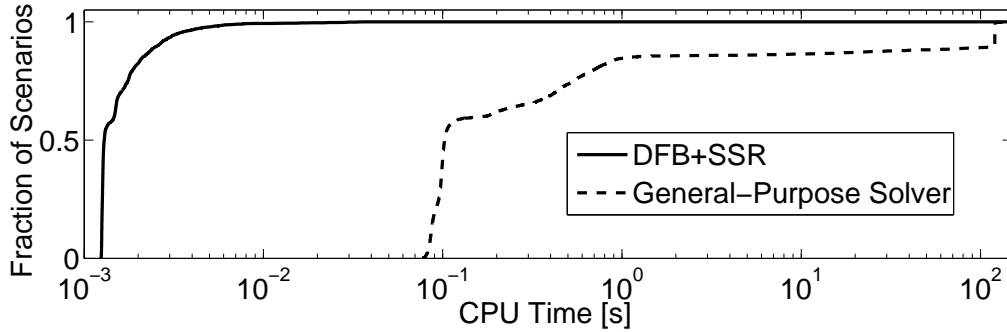
Figure 3.4: CPU time comparison between our two proposed search mechanisms (DBF and SSR) and an applicable general-purpose solver under a time-out of 2 minutes, in randomly sampled 5-user ADSL2 scenarios.

parts, where in the first one in Lines 4–8 we solve the lower-bounding problem in (3.4) as described in Appendix A.7 and search a feasible target objective value L^{target} for the exact problem in (3.2). Note that also the objective value of the current *global* incumbent solution can be used as the target objective value. While the incumbent objective might be lower than any feasible objective value in the subtree considered in the exact problem, the SSR scheme will not exclude any solution in the subtree from the search-space that has a better objective than the current incumbent and therefore supports the pruning of suboptimal solutions. The second part of the SSR scheme is detailed in Lines 10–14 of Algorithm 12 and concerns the search-space reduction. As seen in Appendix A.7 the lower-bounding problem can be solved by separate problems for users $i, u+1 \leq i \leq U$. Hence, also the minimum and maximum bit-rate of user i in the subtree rooted at $\mathbf{r}^{(u)}$ can be found by searching, starting from the optimum of the lower-bounding problem, the minimum and maximum rates which result in a lower-bound that is below the target objective value.

This SSR method can be employed at any level of the search-tree of BFB and DFB, cf. Line 16 in Algorithm 11, but only on the root node in RSB. Note also that minimum power/maximum bit-loading information is obtained by SSR which is used for computing lower-bounds in Line 18 of Algorithm 11.

3.3 Simulation Results on Optimal DSM

3.3.1 Performance comparison to a general-purpose solver

In order to motivate our problem-specific BnB and variable-range reduction mechanisms DFB and SSR we compare them to the open-source general-purpose solver for non-convex mixed-integer problems "Couenne" [8] in terms of central processing

unit (CPU) time for various sum-rate maximization problems.[2] In order to have problem instances that are easier to solve by general purpose software we base the simulations on 5-user upstream transmission ADSL2 scenarios[3], with topologies generated by uniformly sampling the users' line-lengths between 500 m and 2000 m. The used crosstalk channel data is based on measurements in [153] and a random cable selection as the number of measured lines exceeds U. Figure 3.4 depicts the cumulative distributions of CPU time for our combined scheme (DFB and SSR) and for Couenne. The main observation from these results is that there is a subset of the per-subcarrier problems which could not be solved by Couenne in the set time-limit of 120 seconds. Hence, the development of problem-specific techniques for the problem at hand is qualitatively justified.

3.3.2 Comparison of BnB methods

In the following we investigate the average complexity of solving the subproblems in (3.1) under the three BnB schemes outlined in Section 3.1. In order to study a more challenging DSM case we use numerous VDSL upstream scenarios as described below and a 99 % worst-case crosstalk model [59]. For tractability we restrict ourselves to the simulation of every 50 th subcarrier out of the more than 1600 subcarriers under fixed values of weights $\breve{w}_u = 0, \hat{w}_u = 1/U$, and identical Lagrange multipliers $\lambda_u = \lambda, \nu_u = 0, \forall u \in \mathcal{U}$. To capture the performance of these methods exclusively we do not make use of incumbent initializations. Also, we avoid machine dependency of our performance evaluation to a large extent by focussing on two reproducible complexity merits: the number of visited feasible leaf nodes (rate allocations) $\mathbf{r} \in \mathcal{Q}^{\mathbf{r}}$, and the total number of visited leaf nodes $\mathbf{r} \in \mathcal{L}$ (corresponding to the number of solved matrix equations [33] in the power evaluation $\mathbf{p}(\mathbf{r})$ for a bit-allocation $\mathbf{r}$), where the latter also includes the tested allocations $\mathbf{r} \in \bar{\mathcal{Q}}^{\mathbf{r}}$ which turn out infeasible. Furthermore, we neglect the complexity of other logical operations needed to perform the BnB searches. As explained in Section 3.1.2 the RSB scheme has more severe memory requirements compared to our two BnB search proposals, and hence our

---

[2]Both algorithms solve, for each scenario, the per-subcarrier problem in (3.1) for all subcarriers $c \in \mathcal{C}$ only once, that is, for the set of weights and Lagrange multipliers that are output in the first iteration by the sum-rate maximizing master LP in (4.1) after being initialized by the solution of the DBL algorithm in [103]. For Couenne we apply default parameters, except for the branching priority of continuous variables which was set lower than for integer variables. The platform is Windows 7 on an Intel quad-core system running at 2.4 GHz with 4 GB of random-access-memory.

[3]The parameters for ADSL2 and VDSL follow the corresponding standards in [86] (Annex B.1.3) and in [50] (band plan 997-M1x-M, a flat spectral mask constraint at -60 dBm/Hz, alien crosstalk according to VDSL noise A [50]), with $\Gamma = 12.8$ dB and $\theta = 1$. The background noise for ADSL2 is set to $N_u^c = -120$ dBm/Hz, while that for VDSL is chosen as -140 dBm/Hz. The maximum bit-allocation $\hat{\theta}$ for the single-carrier simulations in Sections 3.3.1–3.3.3 was set to 16 as it was used in the BnB simulations in [163], while for multi-carrier simulations we use a standard compliant value of 15.
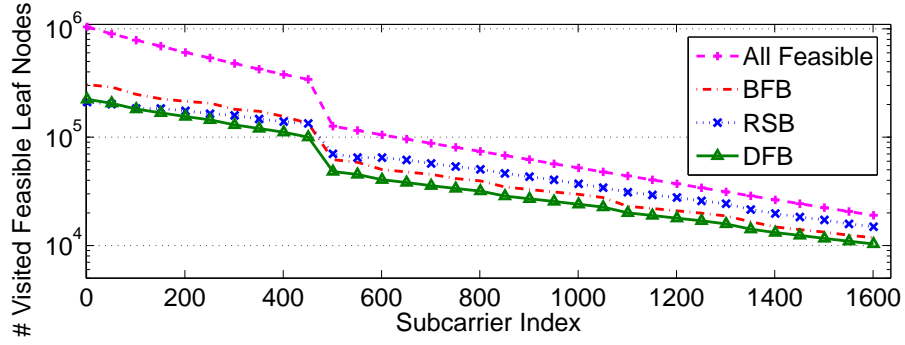
(a)
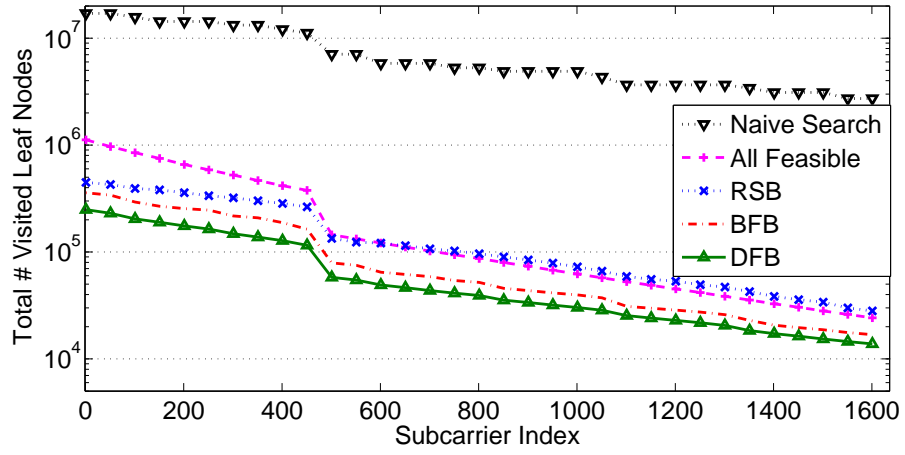
(b)
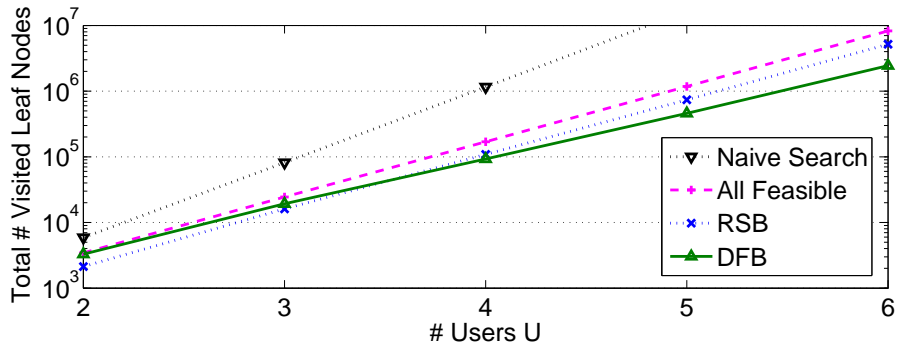
(c)

Figure 3.5: Average complexity of various BnB schemes in terms of the number of total or feasible visited leaf nodes for solving the per-subcarrier problems in (1.12) with $\hat{w}_u = 1/U, \lambda_u = 10^{-3}, \forall u \in \mathcal{U}$; in a) and b) over the subcarrier index for $U = 6$; in c) the sum-complexity for all per-subcarrier problems over the number of users U.

comparisons favor this previously proposed scheme. The network scenarios are based on a set of specified line lengths $\{200, 400, 600, 800\}$ m, and forming all U-combinations with repetitions[4] to allocate users to these lengths, cf. [181] for further details.

Figures 3.5(a) and 3.5(b) show the total number of visited leaf nodes and the number of visited feasible leaf nodes, respectively, over the subcarrier index under different search schemes for 6 users. The bend of the curves at index 500 is due to the use of two non-adjacent frequency sub-bands. In general we find that the complexity in solving the per-subcarrier problems (3.1) decreases in all schemes with the subcarrier frequency due to the increasing crosstalk coupling per unit-length and channel attenuation. In a "naive search" all allocations in $\prod_{u \in \mathcal{U}} \{0, \theta, \ldots, \hat{\theta}_u\} \subseteq \mathcal{L}$ are evaluated, where we define $\hat{\theta}_u, \hat{\theta}_u \leq \hat{\theta}$, as the maximum number of bits the user u can transmit without interference, cf. Figure 3.5(b). However, we see that the number of feasible leaf nodes $|\mathcal{Q}^{\mathbf{r}}|$ in the search-tree, labelled "all feasible", is far below this number. This complexity reduction can already be achieved by a modified exhaustive search [188, Algorithm 2]. Furthermore, we observe a reduction in the number of visited leaf nodes by the BnB schemes BFB, DFB and RSB, with DFB performing best over the whole range of subcarrier problems, cf. Figure 3.5(b). This can be further explained by comparing Figures 3.5(a) and 3.5(b), where we see that a large part of the complexity of RSB lies in the evaluation of *infeasible* allocations, which is avoided by our proposed BnB mechanism in Section 3.1.2.1. Regarding the comparison in average sum-complexity over subcarriers in Figure 3.5(c) we see that for less than 4 users RSB performs better than DFB. However, for a higher number of users we observe a growing gap in complexity between the two schemes with DFB performing better, both, in terms of computational and memory complexity.

3.3.3 The impact of the Lagrange multipliers on the complexity

We use the same simulation setup as in the previous section and demonstrate in Figure 3.6 the dependency of the average (over users and network scenarios) search complexity on the Lagrange multipliers $\lambda_u = \lambda, u \in \mathcal{U}$, associated with the bit-rates. Our SSR scheme is shown to successfully reduce the search space for lower values of λ, where we note that the algorithm is only executed once at the root of the search-tree, cf. the curve labelled "DFB+SSR, No Bounding". Higher values of λ on the other hand lead to higher optimal per-subcarrier rates and therefore to stronger interference levels among users at the optimum. This on the other hand decreases the quality of the lower-bound computed from the interference-free problem in (3.4) and therefore the performance of the SSR scheme. However, DFB outperforms RSB for larger values of λ, making the joint application of SSR and DFB search outperform RSB over the whole range of λ, cf. the curve labelled "DFB+SSR, With Bound-

---

[4]For example, for $U = 6$ the number of thereby generated scenarios is 84.
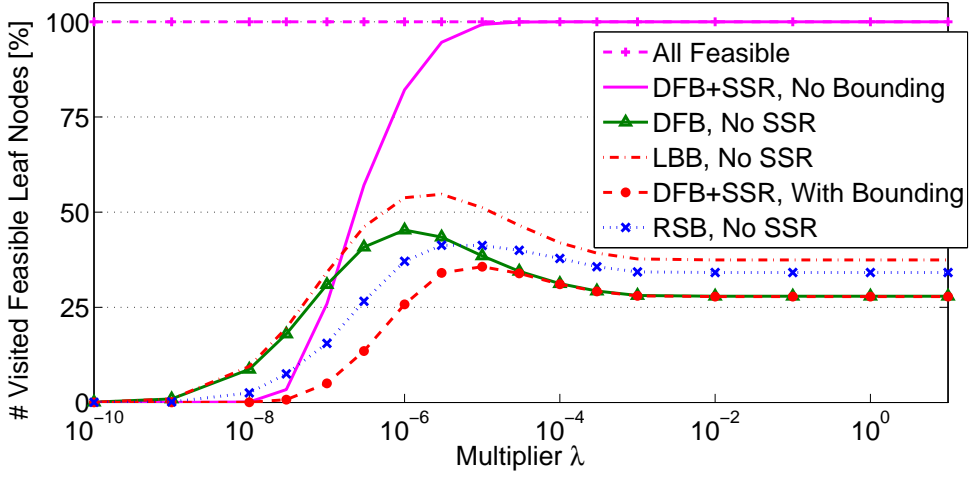
(a)
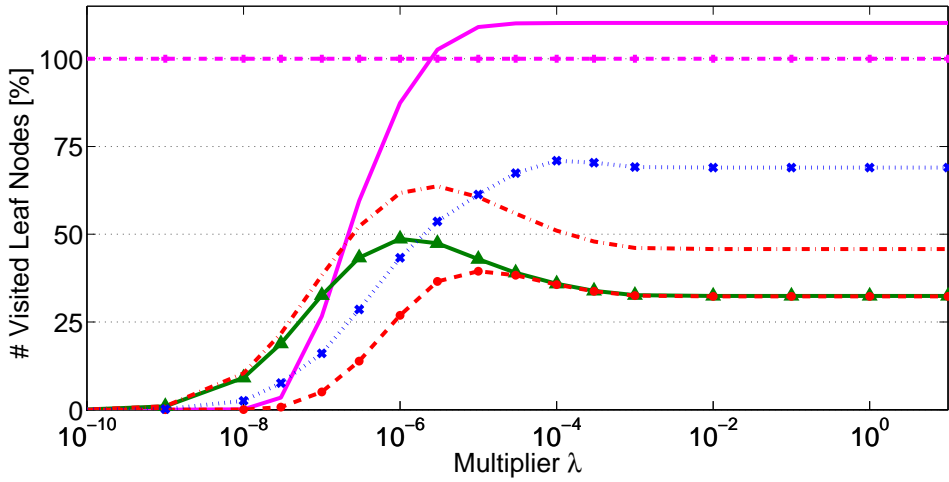
(b)

Figure 3.6: Dependency of the average branch-and-bound (BnB) sum-complexity over all selected per-subcarrier problems in (1.12) on the Lagrange multiplier $\lambda_u = \lambda, \forall u \in \mathcal{U}$, in various 6-user VDSL scenarios; a) Visited feasible bit-allocations; b) Total number of tested bit-allocations.

ing". By comparison between Figures 3.6(a) and 3.6(b) we see once more that our proposed BnB schemes avoid testing infeasible bit-allocations, while a large part of the complexity of RSB lies in testing such allocations. Furthermore, as explained in Section 3.1, RSB has a worst-case storage requirement that increases exponentially
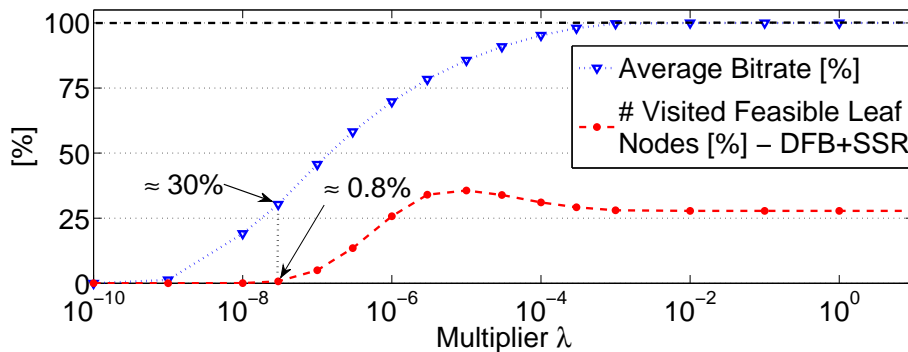
Figure 3.7: Comparison between the average sum-complexity and the achieved per-user bit-rates for our proposed search scheme.

in the number of users, while that of DFB increases only linearly. Altogether we argue that DFB is the preferable BnB scheme, especially in scenarios with a larger number of users.

Another aspect in Figure 3.6 is that the average search complexity of all shown BnB schemes only increases up to a certain value of λ, after which it decreases again. As the number of feasible allocations is independent of λ we attribute this behavior to the lower-bound $lb^{\mathbf{r}^{(u)}}(\mathbf{p}^{\min}, \mathbf{r}^{\max})$ in (3.3). We argue that during BnB search there is typically a good bound $\mathbf{r}^{\max}$ available, be it either based on the search-region definition as in RSB [163] or on message passing of maximum bit-loading information as in DFB and BFB. However, lower-bounds $p_i^{\min}$ on the optimal power consumption of a user $i, u < i \leq U$, for a sub-tree rooted at $\mathbf{r}^{(u)}$ are harder obtained, where notably our SSR scheme allows us to get lower-bounds other than $p_i^{\min} = 0$, cf. Line 14 of Algorithm 12. Altogether we have stronger lower-bounds for either lower values of λ where the optimal power levels are small, or higher values of λ where the total weight $(\check{\mathbf{w}} + \boldsymbol{\lambda})$ on the rates is dominant.

The multiplier λ controls the rates' weight in the objective $f(\mathbf{p}, \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda})$ defined in (1.9). Figure 3.7 illustrates the average per-user rate over the multiplier λ in percent of the maximum achieved value (at $\lambda = 10$) and compares it to the complexity of our combined search method (DFB and SSR) taken from Figure 3.6(a). We observe that a wide range of low average rates is achievable at fairly low complexity. For example, we find that for $\lambda = 3 \cdot 10^{-8}$ our combined search method (DFB and SSR) visits less than 0.8% of the total number of feasible leaf nodes in the search-tree, while the resulting average rate is more than 30% of the maximum one at $\lambda = 10$.

3.3.4 The impact of target-rates on the complexity of DSM

Differently to the previous section we will now directly analyze the impact of the target-rates on the complexity of a dual optimal DSM scheme as a whole in
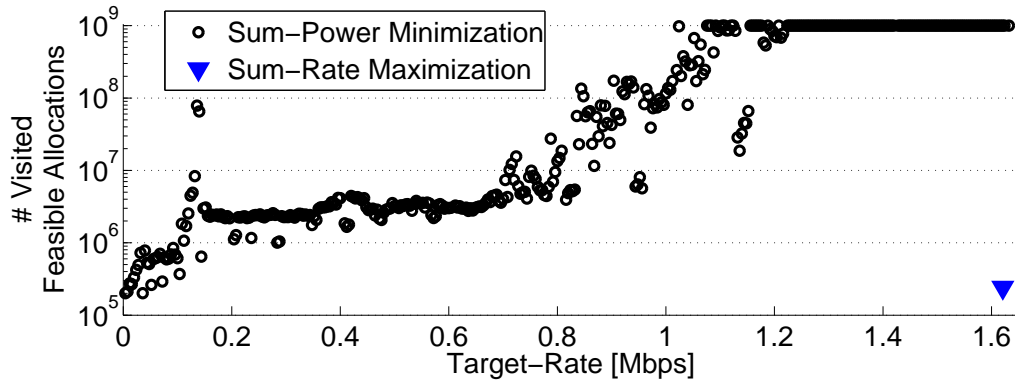
Figure 3.8: Complexity of solving a dual sum-power minimization problem using the column generation technique of Chapter 4 and the optimal techniques DFB and SSR in a distributed 16-user ADSL2 scenario.

various sum-power minimization problems. The simulation is based on a single ADSL2 topology with 16 users transmitting in upstream direction and situated at $800, 850, \ldots, 1550\,\mathrm{m}$ distance from the deployment point, and we use the average crosstalk per unit-length between all 24 measured cables in [153] for the calculation of the crosstalk couplings. A maximum total number of evaluated feasible allocations of 10^9 was set to limit the simulation time. From the simulation results shown in Figure 3.8 we find that the complexity of our dual-optimal DSM algorithm qualitatively increases with the target-rates. However, we also note that sum-rate maximization (cf. the point at the bottom-right in Figure 3.8) had again a complexity which was lower compared that of sum-power minimization at the highest shown target-rates. This is most likely due to the fact that sum-power constraints were not tight in this scenario and the multipliers $\boldsymbol{\nu}$ are consequently all zero. Therefore a good lower bound on the objective can be computed by knowing a good upper-bound on the number of bit-steps that can be loaded, as obtained in our DFB search, cf. Section 3.1.2.1. This observation is also in accordance with the interpretation of Figure 3.6(a) in Section 3.3.3.

3.3.5 Performance evaluation of greedy multi-user bit-loading

We will next investigate the suboptimality in rate-maximization problems of two multi-user discrete bit-loading (DBL) approaches: the greedy multi-user DBL (MDBL) scheme in [103] and single-user bit-loading [27] under worst-case crosstalk noise computed using the users' PSD masks and the (measured) cross-channel data (i.e., "Mask-Based" SSM). The performance metric we consider for these two heuris-
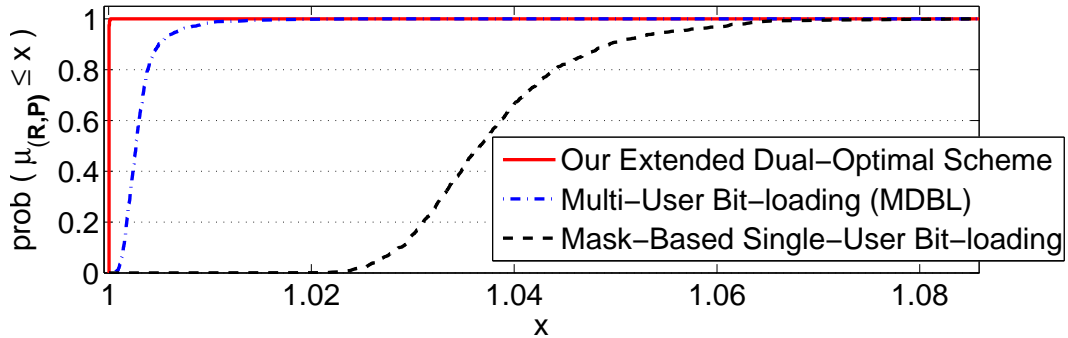
Figure 3.9: Cumulative distribution function of the performance ratio $\mu_{(\mathbf{R},\hat{\mathbf{P}})}$ in (3.5) for our dual-optimal DSM scheme and two bit-loading algorithms.

tics is their performance ratio defined as

$$\mu_{(\mathbf{R},\hat{\mathbf{P}})} = \frac{D^*_{(\mathbf{R},\hat{\mathbf{P}})}}{P^{\text{heur}}_{(\mathbf{R},\hat{\mathbf{P}})}}, \tag{3.5}$$

where we denote the objective value of a heuristic bit-loading scheme achieved by a feasible solution of the primal problem in (1.8) as $P^{\text{heur}}_{(\mathbf{R},\hat{\mathbf{P}})} \geq P^*_{(\mathbf{R},\hat{\mathbf{P}})}$. For example, in the considered case of $\hat{\mathbf{w}} = \mathbf{0}$, meaning a pure rate-maximization problem, we have $D^*_{(\mathbf{R},\hat{\mathbf{P}})} \leq P^{\text{heur}}_{(\mathbf{R},\hat{\mathbf{P}})} \leq 0$ and it therefore holds that $\mu_{(\mathbf{R},\hat{\mathbf{P}})} \geq 1$. The dual optimal objective $D^*_{(\mathbf{R},\hat{\mathbf{P}})}$ in (1.10) is computed by our dual optimal DSM scheme, consisting of the optimal Lagrange multiplier update algorithm in Chapter 4, initialized at the solution under MDBL, and using our optimal techniques DFB and SSR for the subproblems in (3.1). In order to obtain a clearer picture of the duality-gap $(P^*_{(\mathbf{R},\hat{\mathbf{P}})} - D^*_{(\mathbf{R},\hat{\mathbf{P}})})$ in realistic scenarios we also derive primal feasible solutions to the problem in (1.8) by applying the heuristic described in Section 4.1.4 on top of our dual optimal DSM scheme. We simulated 1000 ADSL2 scenarios with 10 users and other simulation parameters and a random topology and cable selection as described in Section 3.3.1. Our results are shown in Figure 3.9. Regarding the cumulative distribution[5] for our extended dual-optimal scheme we see that the performance ratio $\mu_{(\mathbf{R},\hat{\mathbf{P}})}$ was in fact nearly 1 in all tested scenarios. More precisely, its suboptimality and therefore the duality-gap is below 0.01% of $D^*_{(\mathbf{R},\hat{\mathbf{P}})}$ in more than 99.6% of the scenarios, with a confidence of 99% according to a t-test. Similarly, the suboptimality of MDBL is guaranteed to lie below 1% in more than 97.7% of the scenarios, demonstrating the near-optimality of MDBL in realistic ADSL2 scenarios. Surprisingly, even the conservative mask-based SSM approach leads to a suboptimality of below 5% in more than 91.1% of the scenarios.

---

[5]In the computer science literature [45] this empirical distribution is referred to as "performance profile".

Chapter 4 Low-Complexity Discrete-Rate DSM

In this chapter we study *low-complexity* algorithms for the discrete, constrained and interference-limited multi-user and multi-carrier power control problem in (1.8). Lagrange relaxation is a common technique for the decomposition of such problems into independently solvable per-subcarrier subproblems. However, in the previous chapter we saw that the optimal solution of these subproblems as demanded by Lagrange relaxation is not tractable to compute in large DSL networks in general.

Spectrum balancing algorithms applicable for a large number of users are currently based on further complexity reduction heuristics such as constructive greedy heuristics [103, 196], continuous and/or (sequential) convex relaxation [66, 160, 197], and sequential power updates over users [31, 160, 197, 198].

An alternative to the Lagrange relaxation is the time-sharing relaxation which allows for a convex combination of various power-allocation solutions. Furthermore, as shown in Section 2.1.2, it is the strong dual problem to the Lagrange relaxation, also when the set of transmission rates is finite. Optimal subcarrier and power-allocation in OFDMA networks with *continuous* power-allocation was recently shown to have a polynomial time approximation [114], where also an equivalent linear time-sharing formulation was derived. Similarly, continuous time-sharing was shown to yield a convex and therefore polynomially solvable optimization problem under continuous power-allocation in [80, 140, 190], cf. Section 2.1.4 for details.

In DSM problems with inter-user interference time-sharing was introduced as a method which schedules various multi-user power-allocations over time, each of these allocations still allowing for inter-user interference [56, 187]. In Section 2.1.2 we analyze the time-sharing relaxation of our primal problem in (1.8). Its intractability motivates the proposal of a novel framework for multi-carrier power control based on a nonlinear Dantzig-Wolfe (NDW) decomposition [44] and a problem "disaggregation" [132] in Section 4.1. This approach differs from previous Lagrange relaxation schemes for DSM in interference-limited systems [33, 52, 163, 198] in the dual master problem which is a linear program giving time-sharing solutions and which exploits independence among subcarriers by separate treatment of the per-subcarrier solutions. We emphasize that the time-sharing solutions do in our case still allow for inter-user interference. In Section 4.1.2 we sketch how NDW decomposition can not only be applied to sum-rate and sum-power optimization but also for the maximization of users' minimum rate, geometric mean rate, harmonic mean

rate, proportional fair rate [151], and weighted minimum rate. Differently to most work on dual-relaxation based DSM schemes, in Section 4.1.4 we also suggest a heuristic to recover a feasible solution to the original primal problem. The proposed DSM method avoids numerical convergence problems arising due to similarity of subcarriers [52] or a positive duality-gap [114]. Furthermore, it bears the potential to use a combination of optimal and low-complexity suboptimal solutions to the combinatorial per-subcarrier subproblems while providing a monotonously improving objective value. This is in contrast to most previous Lagrange-relaxation based DSM algorithms where the dual master problem theoretically demands for optimal subproblem solutions, cf. (1.12), an exception being for instance the scheme in [111] which uses approximate subgradients [12].

While we have seen in Section 2.1.2 that the combinatorial per-subcarrier problems in (1.12) have polynomial complexity in the number of users and developed low-complexity optimal schemes in Chapter 3, we find that such suboptimal methods are indispensable in large systems. Thus we give an overview of various basic dual heuristics, that is greedy and local search, and randomized extensions thereof, including a "warm-start" heuristic which exploits the correlation among the solutions of the per-subcarrier problems on neighboring subcarriers. The fact that the number of heavily disturbing lines in near-far DSL scenarios is limited was exploited in proposed schemes for vectored DSL systems by limiting the number of canceled crosstalk sources [32]. We propose a mixed exhaustive and greedy dual heuristic which makes use of this network feature to reduce the complexity also for spectrum balancing. The simulation results following in Section 4.3 demonstrate the advantage of performing a heuristic combinatorial search jointly for all users. Besides a large 50-user VDSL network example we also provide an average DSM performance comparison in a large set of thousand VDSL scenarios with mixed central office (CO) and cabinet deployment.

4.1 A Novel Framework for DSM

In this section we propose a novel DSM framework which approaches the original problem in (1.8) by iteratively optimizing the time-sharing relaxation in (2.3) and using a heuristic for recovering feasible solutions for the original problem. Its key features are the decomposition into independent per-subcarrier problems, similar to the Lagrange dual problem in (1.10), as shown in Section 4.1.1, its applicability to various DSM objectives as highlighted in Section 4.1.2, and the possibility for heuristic solutions of the per-subcarrier problems as studied in Section 4.2. Our method is related to previous dual relaxation based DSM algorithms [33, 52, 163, 198] through Theorem 1 and partly motivated by the results in [114, 198] and Section 2.2, showing a vanishingly small duality-gap — that is the difference between the optimal objectives in the original problem in (1.8) and its dual in (1.10). The time-sharing

problem in (2.3) can not only be motivated by its connections with Lagrange relaxation, but can also be regarded as the linear relaxation of an otherwise identical problem to that in (2.3) with binary variables $\xi_i^c \in \{0, 1\}, i \in \mathcal{I}_c, c \in \mathcal{C}$. This binary problem [182] only picks a single allocation per subcarrier and is therefore equivalent to the original problem in (1.8). In Chapter 5 we will study an extended binary problem which we partly solve by a semidefinite relaxation [102]. We note however that this relaxation leads to a nearly quadratic increase in the number of variables and we therefore deem its complexity too high for current solvers [155], even when the iterative scheme in the following section is applied. Furthermore, we will propose a heuristic for recovering binary solutions in Section 4.1.4 which shows good performance in simulations (cf. its explicit performance evaluation in Sections 3.3.5 and 4.3.1), making the application of more complex relaxations than the linear relaxation unattractive.

4.1.1 Nonlinear Dantzig-Wolfe Decomposition

The decomposition scheme described next from first principles is based on the mathematical programming concept of column[1] generation [110], or more precisely a nonlinear Dantzig-Wolfe (NDW) decomposition [132], [44, Ch. 23] of (1.8). Applications of this decomposition approach in the area of wireless communication can be found in [6, 19, 91, 148]. At iteration k of the algorithm we consider a subset of all columns $\mathcal{I}^{c,(k)} \subseteq \mathcal{I}^c, c \in \mathcal{C}$ in (2.3), yielding the restricted master problem

$$P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,ts(k)} = \underset{\xi' \geq 0, \xi_i^c \geq 0, i \in \mathcal{I}^{c,(k)}, c \in \mathcal{C}}{\text{minimize}} \quad \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}^{c,(k)}} f^c(\mathbf{p}^{c,i}, \hat{\mathbf{w}}, \breve{\mathbf{w}})\xi_i^c + f'(\hat{\mathbf{P}}, \mathbf{R})\xi' \quad (4.1a)$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}^{c,(k)}} \mathbf{r}^c(\mathbf{p}^{c,i})\xi_i^c + \mathbf{R}\xi' \succeq \mathbf{R}, \quad (4.1b)$$

$$\sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}^{c,(k)}} \mathbf{p}^{c,i}\xi_i^c + \hat{\mathbf{P}}\xi' \preceq \hat{\mathbf{P}}, \quad (4.1c)$$

$$\sum_{i \in \mathcal{I}^{c,(k)}} \xi_i^c + \xi' = 1, \quad \forall c \in \mathcal{C}, \quad (4.1d)$$

where we added an artificial column with weight ξ' and cost $f'(\hat{\mathbf{P}}, \mathbf{R}) = \hat{\mathbf{w}}^T(\hat{\mathbf{P}} + \boldsymbol{\delta}) - \breve{\mathbf{w}}^T(\mathbf{R} - \boldsymbol{\delta})$, for some arbitrary $\boldsymbol{\delta} \succ \mathbf{0}$. By setting $\xi' = 1$ and $\xi_i^c = 0$, for all $i \in \mathcal{I}^{c,(k)}$ and $c \in \mathcal{C}$, it can be seen that this "aggregated" column makes (4.1) always feasible. Furthermore, the choice of cost leads to the following results, indicating that this artificial column does not alter the solution when (2.3) is feasible.

Theorem 8. *At the optimum of (4.1) we have $\xi' = 0$ if a feasible solution to*

---

[1]The term "column" refers to the column-vectors $\mathbf{p}_c^i$ and $\mathbf{r}_c(\mathbf{p}_c^i)$ in the constraint matrices of the LP in (2.3).

(2.3) exists with $\xi_i^c = 0$, for all $i \in \mathcal{I}^c \setminus \mathcal{I}^{c,(k)}, c \in \mathcal{C}$, and $\xi' = 1, \xi_i^c = 0$, for all $i \in \mathcal{I}^{c,(k)}, c \in \mathcal{C}$, otherwise.

See Appendix A.8 for a proof.

Corollary 4. *Assuming feasibility of (2.3) we have $P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,ts(k)} \geq P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,ts}$.*

Proof. By Theorem 8 we either have $\xi' = 0$ or $\xi' = 1$ at the optimum of (4.1). The corollary follows as $\mathcal{I}^{c,(k)} \subseteq \mathcal{I}^c, \forall c \in \mathcal{C}$, and by $\boldsymbol{\delta} \succ \mathbf{0}$ and feasibility in (2.3) any solution of (2.3) has a lower objective than $f'(\hat{\mathbf{P}}, \mathbf{R})$. $\qquad\square$

After solving (4.1), the second task at each iteration in a column generation scheme is to compute new columns to be added to the master problem in (4.1) in order to reduce the gap described in Corollary 4. Relaxing constraints (4.1b), (4.1c) and (4.1d) in the restricted master problem at iteration k, and denoting their Lagrange multipliers by $\boldsymbol{\lambda}^{(k)}, \boldsymbol{\nu}^{(k)} \in \mathbb{R}_+^U$ and $\boldsymbol{\Phi}^{(k)} \in \mathbb{R}^C$, respectively, and also including variables $\xi_i^c = 0$, for all $i \in \mathcal{I}^c \setminus \mathcal{I}^{c,(k)}$ and $c \in \mathcal{C}$, we can write the Lagrangian for (4.1) as

$$L^{(k)} = \sum_{c\in\mathcal{C}} \sum_{i\in\mathcal{I}^c} \left(f^c(\mathbf{p}^{c,i}, \hat{\mathbf{w}} + \boldsymbol{\nu}^{(k)}, \breve{\mathbf{w}} + \boldsymbol{\lambda}^{(k)}) + \Phi_c^{(k)} \right) \xi_i^c +$$
$$\left(\mathbf{R}^T \boldsymbol{\lambda}^{(k)} - \hat{\mathbf{P}}^T \boldsymbol{\nu}^{(k)} - \sum_{c\in\mathcal{C}} \Phi_c^{(k)} \right)(1 - \tilde{\xi}) + \tilde{f}(\hat{\mathbf{P}}, \mathbf{R})\tilde{\xi}. \qquad (4.2)$$

Adding any column $i \in \mathcal{I}^c \setminus \mathcal{I}^{c,(k)}$ to (4.1) with negative derivative[2] $\partial L^{(k)}/\partial \xi_i^c = f^c(\mathbf{p}^{c,i}, \hat{\mathbf{w}} + \boldsymbol{\nu}^{(k)}, \breve{\mathbf{w}} + \boldsymbol{\lambda}^{(k)}) + \Phi_c^{(k)}$ at the (dual) optimum of (4.1) lowers the optimal objective of (4.1) or leaves it unchanged.[3] Hence, a simple criterion, leading in fact to pricing problems similar to those in the Lagrange-dual of the time-sharing problem in (A.2), is to pick the column on subcarrier c with minimal derivative, leading to decomposable subproblems similar to (1.12) (and with identical complexity, cf. Corollary 1) in the form of

$$q^{c,\text{red}}(\boldsymbol{\lambda}^{(k)}, \boldsymbol{\nu}^{(k)}, \Phi_c^{(k)}) = q^c(\boldsymbol{\lambda}^{(k)}, \boldsymbol{\nu}^{(k)}) + \Phi_c^{(k)}, \qquad (4.3a)$$
$$= \min_{i\in\mathcal{I}^c} \left\{ f^c(\mathbf{p}^{c,i}, \hat{\mathbf{w}} + \boldsymbol{\nu}^{(k)}, \breve{\mathbf{w}} + \boldsymbol{\lambda}^{(k)}) \right\} + \Phi_c^{(k)}, \forall c \in \mathcal{C}. \quad (4.3b)$$

The proposed DSM algorithm iterates between solving (4.1) and the C subproblems in (4.3b). Hence, the number of columns in (4.1) increases by at most C in each

---

[2]This derivative is also referred to as the "reduced cost" of a column [14].

[3]More precisely, assuming non-degeneracy of a basic solution of (4.1) one can pivot on the new variable ξ_i^c with $\partial L^{(k)}/\partial \xi_i^c < 0$ and thereby maintain feasibility while *strictly* decreasing the objective value [14]. From a dual perspective, non-degeneracy corresponds to uniqueness of the dual solution $\boldsymbol{\lambda}^{(k)}, \boldsymbol{\nu}^{(k)}, \Phi_c^{(k)}, c \in \mathcal{C}$, to (4.1) [14]. Under this uniqueness it follows from the existence of a negative gradient direction w.r.t. (4.2) at $\boldsymbol{\lambda}^{(k)}, \boldsymbol{\nu}^{(k)}, \Phi_c^{(k)}, c \in \mathcal{C}$, and strict duality that the optimal objective of (4.1) decreases *strictly*.

iteration, cf. Algorithm 1. We emphasize that any potentially suboptimal solution to (4.3b) with negative objective may improve the restricted master problem (4.1), which is the reason why problem (4.3b) is amenable for fast heuristics, cf. Line 7 in Algorithm 1 and our overview on basic heuristics in Section 4.2. Furthermore, we have that if $q^c(\boldsymbol{\lambda}^{(k)}, \boldsymbol{\nu}^{(k)}) \geq -\Phi_c^{(k)}, \forall c \in \mathcal{C}$, then $P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,ts(k)} = P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,\mathrm{ts}}$, cf. (4.3a). The same conclusion can be drawn if, $\forall c \in \mathcal{C}$, we have $i \in \mathcal{I}^{c,(k)}$, where i is the minimum argument in (4.3b), i.e., all newly generated columns are already part of the master LP in (4.1). This follows from $P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,\mathrm{ts}(k)} \geq P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,\mathrm{ts}}$, cf. Corollary 4, and the relations

$$P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,\mathrm{ts}} \geq q(\boldsymbol{\lambda}^{(k)}, \boldsymbol{\nu}^{(k)}) \tag{4.4a}$$

$$= \min_{\{i_c \in \mathcal{I}^c, c \in \mathcal{C}\}} \left\{ \sum_{c \in \mathcal{C}} f^c(\mathbf{p}^{c,i_c}, \hat{\mathbf{w}} + \boldsymbol{\nu}^{(k)}, \breve{\mathbf{w}} + \boldsymbol{\lambda}^{(k)}) + \boldsymbol{\lambda}^T \mathbf{R} - \boldsymbol{\nu}^T \hat{\mathbf{P}} \right\} \tag{4.4b}$$

$$= \min_{\{i_c \in \mathcal{I}^{c,(k)}, c \in \mathcal{C}\}} \left\{ \sum_{c \in \mathcal{C}} f^c(\mathbf{p}^{c,i_c}, \hat{\mathbf{w}} + \boldsymbol{\nu}^{(k)}, \breve{\mathbf{w}} + \boldsymbol{\lambda}^{(k)}) + \boldsymbol{\lambda}^T \mathbf{R} - \boldsymbol{\nu}^T \hat{\mathbf{P}} \right\} \tag{4.4c}$$

$$= P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,ts(k)}, \tag{4.4d}$$

where (4.4a) follows from Theorem 1 and the definition of the dual problem in (1.10), (4.4b) holds by definition, (4.4c) holds due to the assumed availability of the optimal column in the restricted master LP in (4.1), and (4.4d) holds as $\boldsymbol{\lambda}^{(k)}, \boldsymbol{\nu}^{(k)}$ were the optimal dual multipliers in (4.1), and due to strong duality in the restricted master LP. This means that the algorithm terminates if not at least one new allocation on any subcarrier is added to the master problem in (4.1). A finite convergence time of the algorithm then follows from the finiteness of $|\mathcal{Q}^c|, c \in \mathcal{C}$. The negativity of $q^{c,\mathrm{red}}(\boldsymbol{\lambda}^{(k)}, \boldsymbol{\nu}^{(k)}, \Phi_c^{(k)})$ as a necessary criterion for an improving column can also be exploited on each subcarrier to reduce the complexity of solving (4.3b), for instance by using $-\Phi_c^{(k)}$ as the initial incumbent objective used for pruning the search tree in branch-and-bound based algorithms, cf. Chapter 3 and specifically Line 19 in Algorithm 11 for an application.

In general, based on Theorem 1 we can bound the (non-negative) gap $P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,ts(k)} - P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,\mathrm{ts}}$ at iteration k by

$$\tilde{\zeta} = P_{(\mathbf{R},\hat{\mathbf{P}})}^{*,\mathrm{ts}(k)} - \max_{1 \leq j \leq k} \{q(\boldsymbol{\lambda}^{(j)}, \boldsymbol{\nu}^{(j)})\}, \tag{4.5}$$

where $q(\cdot, \cdot)$ is defined in (1.11), cf. Line 6 of Algorithm 1. Our stopping criteria in Line 2 of Algorithm 1 additionally include a primal improvement criterion which is necessary when suboptimal solutions are used as proposed in Line 7. Another practical measure is taken in Line 8 where we only add *new* columns to the NDW master problem in (4.1) to reduce its size.

Algorithm 1 NDW-DSM

1: Initialize $k = 1$, $\mathcal{I}^{c,(1)}, \forall c \in \mathcal{C}, \tilde{\zeta} = \infty$, $P^{*,ts(0)}_{(\mathbf{R},\hat{\mathbf{P}})} = \infty$, $\bar{\delta}$

2: **while** $\tilde{\zeta} > \zeta^{\text{tgt}}$ and $|P^{*,ts(k-1)}_{(\mathbf{R},\hat{\mathbf{P}})} - P^{*,ts(k)}_{(\mathbf{R},\hat{\mathbf{P}})}| > \bar{\delta} |P^{*,ts(k-1)}_{(\mathbf{R},\hat{\mathbf{P}})}|$ **do**

3:     Solve (4.1) by a primal-dual LP solver, obtaining dual multipliers $[\boldsymbol{\lambda}^{(k)}, \boldsymbol{\nu}^{(k)}, \boldsymbol{\Phi}^{(k)}]$

4:     Obtain new allocations solving (4.3b) either ...

5:        ... a) Optimally (e.g., by branch-and-bound as in Chapter 3)

6:           Compute $\tilde{\zeta}$ as in (4.5)

7:        ... b) Suboptimally (initially, e.g., using Algorithm 4)

8:     Add only new allocations $i \in \mathcal{I}^c \setminus \mathcal{I}^{c,(k)}$ to (4.1), $k = k + 1$

9: **end while**

10: Apply Algorithm 2 to recover solutions to (1.8)

4.1.2 Further Properties of NDW-DSM

We claim that the presented disaggregated NDW-DSM algorithm is numerically more stable than previous Lagrange relaxation based DSM schemes based on the following observations: As it is based on a time-sharing formulation it does not suffer from the convergence problems [52,114] which may arise due to non-convexity of the original problem in (1.8). In general, computing a primal feasible allocation based on the dual optimum is then again non-deterministic polynomial-time (NP) hard [114]. However, by Theorem 1 there exists a time-sharing solution having an objective value equal to the optimal dual one. In [52,53] a specific Lagrange multiplier search scheme was proposed which similarly to our scheme yields time-shared solutions but works differently with an aggregated formulation and theoretically necessitates optimal per-subcarrier allocation schemes. We point out that the presented basic NDW-DSM scheme may be further improved by stabilization techniques, cf. [110] for an overview. We have already pointed out several times that NDW-DSM allows for sub-optimal per-subcarrier bit and power-allocation procedures for the discrete, non-convex per-subcarrier problems in (4.3b), while providing a monotonously improving objective in (4.1a) over iterations k. Furthermore, the NDW master problem in (4.1) can be initialized with the solution obtained from suboptimal algorithms such as iterative spectrum balancing [31,198] assuming the per-subcarrier feasibility in (1.7) is met, thereby extending previous DSM schemes. This initialization may happen either by initializing the set $\mathcal{I}^{c,(1)}, c \in \mathcal{C}$, using the per-subcarrier solutions, or by adding the sum-rate and sum-power solution in a similar way we added the artificial column in the NDW master problem in (4.1). The memory of per-subcarrier solutions in the NDW master problem in (4.1) may also pay off in cases where the original problem in (1.8) needs to be solved for various values of target-rates $\mathbf{R}$ or objective functions, see Section 4.1.3 and Chapter 5 for examples.

4.1.3 Alternative Optimization Objectives

We have shown how the NDW decomposition can be applied to the optimization of sum-rate and sum-power as covered by the objective in (1.9). More generally, any minimization of a convex objective of users' sum-power or sum-rates can be approached by the NDW decomposition, yielding a convex master problem eventually including auxiliary sum-rate or sum-power variables and decomposable subproblems in the form of (4.3b). Examples include the maximization of the users' minimum rate, the "balanced capacity" [93] (which is nothing but a weighted minimum rate maximization problem, cf. the similar projection problem in (4.6) below), the concave [22, Sec. 3.1.5] users' geometric mean rate, the concave [22, Ex. 3.17] users' harmonic mean rate [81], or the concave "α-fairness" utility family [151], with weighted proportional sum-rate fairness being one example of this family. To see this consider the latter objective given by [151] $\sum_{u \in \mathcal{U}} w_u \log(t_u)$, where t_u are auxiliary sum-rate variables which constrain the sum-rate as $\sum_{c \in \mathcal{C}} \mathbf{r}^c(\mathbf{p}^c) \succeq \mathbf{t}$ similarly to the target-rates in (1.8b). It can readily be verified following the same steps as in Section 4.1.1 that NDW decomposition results in a convex master problem over the same variables as in (4.1) and additionally the U auxiliary variables $\mathbf{t}$, as well as decomposable subproblems identical in form to those in (4.3b).

Yet another application of our framework can be found in the context of optimization across protocol layers [24]. More precisely, by the approach in [24] any (e.g., non-concave) utility function can be locally optimized by an intertwined sequence of utility-gradient based updates of target-rates $\mathbf{t}$ and projections of these target-rates onto the rate-region (e.g., the set of users' sum-rates achievable by our DSM algorithm) in a direction $\bar{\mathbf{n}} \in \mathbb{R}^U$. We refer to [24] for details on this iterative projected-gradient decomposition scheme. However, the most work-intensive task in this scheme is the projection step, which, adapted to our notation and setting in DSL, can be written as the "weighted" minimum-rate maximization problem

$$\underset{\mathbf{p}^c \in \mathcal{Q}^c, c \in \mathcal{C}, x}{\text{maximize}} \quad x \tag{4.6a}$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} r_u^c(\mathbf{p}^c) \ \geq \ t_u + x \cdot \bar{n}_u, \quad \forall u \in \mathcal{U}, \tag{4.6b}$$

$$\sum_{c \in \mathcal{C}} p_u^c \ \leq \ \hat{P}_u, \quad \forall u \in \mathcal{U}. \tag{4.6c}$$

The applicability of the NDW decomposition to this projection problem follows by similar arguments as for the other objectives above, where the master problem remains an LP and the subproblems are once more identical in form to those in (4.3b). The scheme in [24] further necessitates the weights $\breve{\mathbf{w}}$ we would need in a weighted sum-rate maximization problem in order to achieve the same sum-rate

as found at optimum of the projection problem in (4.6).[4] However, we recognize that these are simply given by the dual multipliers associated with the constraints in (4.6b). As time-sharing lets us work on the convex hull of the rate-region, an attractive option is to always optimize over this convex rate-set and apply the last step in Line 10 of Algorithm 1 only in the last iteration of the scheme in [24]. A second advantage of the NDW decomposition in the context of the projected-gradient decomposition in [24] is again the memory of various previously found power-allocations. Intuitively, as the iterates of the target-rates $\mathbf{t}$ can be expected to be reasonably close to each other, the previously found allocations give a good indication for the optimal projected rates and dual variables in the following iterates.

Another application example of our NDW-DSM framework can be found in the context of energy-efficient networks [118], considering the minimization of the ratio $f^{eff}(\mathbf{s}, \mathbf{t}) = (\sum_{u \in \mathcal{U}} s_u / (\sum_{u \in \mathcal{U}} t_u)$ [Joule/Bit], where $\mathbf{s}, \mathbf{t} \in \mathbb{R}^U$ are artificial variables representing the sum-powers and sum-rates of all users, respectively. A similar power-efficiency metric has been suggested for fixed broadband equipment in [51], which is given by the ratio of the average power consumption over the product of bit-rate and loop length. As $f^{eff}(\mathbf{s}, \mathbf{t})$ is a quasi-linear function [22, Ex. 3.32], one general approach [22, Sec. 4.2.5] to minimize this function is by a line-search for the smallest constant K for which the following problem is feasible

$$\underset{\mathbf{p}^c \in \mathcal{Q}^c, c \in \mathcal{C}, \mathbf{s} \succeq \mathbf{0}, \mathbf{t} \succeq \mathbf{0}}{\text{minimize}} \quad 0 \tag{4.7a}$$

$$\text{subject to} \quad K \sum_{u \in \mathcal{U}} t_u \geq \sum_{u \in \mathcal{U}} s_u, \tag{4.7b}$$

$$\sum_{c \in \mathcal{C}} r_u^c(\mathbf{p}^c) \geq \max\{t_u, R_u\}, \quad \forall u \in \mathcal{U}, \tag{4.7c}$$

$$\sum_{c \in \mathcal{C}} p_u^c \leq \min\{s_u, \hat{P}_u\}, \quad \forall u \in \mathcal{U}. \tag{4.7d}$$

By similar arguments as above it follows that after NDW decomposition these problems necessitate the solution of a linear master problem and subproblems as in (4.3b). Similar as in the gradient projection scheme above, the memory of power-allocations found during optimization can be reused when the problem in (4.7) is resolved for different values of K.

4.1.4 Combination Heuristic for Time-Sharing (CHET)

Algorithm 1 targets the solution of the time-shared problem in (2.3), having the same optimal objective as the dual problem (1.10). In most of the previous work on

---

[4]In [24] the projection problem in (4.6) is solved by Lagrange relaxation and the resulting master problem by a subgradient search, and similarly by a cutting plane method in [25]. Hence, these weights are simply the dual optimal Lagrange multipliers found by these methods.

Algorithm 2 Combination Heuristic for Time-sharing (CHET)

1: $[\{i_c^*\}_{c\in\mathcal{C}}] = \text{CHET}(\{\mathcal{I}^{c,(k)}, \xi_i^c, \forall i \in \mathcal{I}^{c,(k)}\}_{c\in\mathcal{C}}, \boldsymbol{\lambda}, \boldsymbol{\nu})$
2: Initialize $\delta, \kappa \geq 0$, $i_c^* = \text{argmax}_{i\in\mathcal{I}^{c,(k)}}\{\xi_i^c\}, \forall c \in \mathcal{C}$
3: **while** No feasible solution to (1.8) found **do** $\delta = \delta * 2$
4:   **while** Allocation i_c^* is updated on any $c \in \mathcal{C}$ **do**
5:    **for** $\forall c \in \mathcal{C}$ **do** $\hat{\mathbf{P}}^* = \sum_{c\in\mathcal{C}} \mathbf{p}^{c,i_c^*}$, $\mathbf{R}^* = \sum_{c\in\mathcal{C}} \mathbf{r}^c(\mathbf{p}^{c,i_c^*})$
6:     $\breve{\mathcal{I}}^c = \{i \in \mathcal{I}^{c,(k)} \mid f^c(\mathbf{p}^{c,i}, \hat{\mathbf{w}} + \boldsymbol{\nu}, \breve{\mathbf{w}} + \boldsymbol{\lambda}) \leq$
      $f^c(\mathbf{p}^{c,i_c^*}, \hat{\mathbf{w}} + \boldsymbol{\nu}, \breve{\mathbf{w}} + \boldsymbol{\lambda}) + (|f^c(\mathbf{p}_c^{i_c^*}, \hat{\mathbf{w}} + \boldsymbol{\nu}, \breve{\mathbf{w}} + \boldsymbol{\lambda})| + \kappa) \cdot \delta\}$
7:     $\Delta R_u^i = [R_u - R_u^* + r_u^c(\mathbf{p}^{c,i_c^*}) - r_u^c(\mathbf{p}^{c,i})]_+, \forall u \in \mathcal{U}, i \in \breve{\mathcal{I}}^c$
8:     $\Delta\hat{P}_u^i = [\hat{P}_u^* - p_u^{c,i_c^*} + p_u^{c,i} - \hat{P}_u]_+, \forall u \in \mathcal{U}, i \in \breve{\mathcal{I}}^c$
9:     $i_c^* = \text{argmin}_{i\in\breve{\mathcal{I}}^c}\{(\hat{\mathbf{w}} + \boldsymbol{\nu})^T \Delta\hat{\boldsymbol{P}}^i + (\breve{\mathbf{w}} + \boldsymbol{\lambda})^T \Delta\boldsymbol{R}^i\}$

---

dual-relaxation based DSM algorithms the problem of recovering feasible solutions for the original problem (1.8) is either overlooked or circumvented by proposing direct approximative implementations of time-sharing solutions [52, 70, 114, 187, 198], with an exception being [52, 53]. However, we found that the heuristics in [52, 53], originally proposed for a specific DSM algorithm and rate-maximization problem, may result in large performance losses when applied to sum-power minimization problems, cf. Section 4.3.1 for an example. More precisely, the scheme in [53] uses the distance to a target (sum-power / sum-rate) solution as the decision metric for greedily selecting a per-subcarrier solution. Furthermore, each user's rate and transmit power are normalized by its target-rate and the maximum sum-power, respectively, which influences the algorithm's valuation of power compared to rate. Hence, the greedy selection heavily depends on this normalization, and only indirectly on the actual objective function. Our novel heuristic proposed in Algorithm 2 remedies this drawback by explicitly taking the optimization objective into account. While after convergence of Algorithm 1 restricted subsets of columns per-subcarrier $\mathcal{I}^{c,(k)} \subseteq \mathcal{I}^c$ are available, enumerating the product set $\prod_{c\in\mathcal{C}} \mathcal{I}^{c,(k)}$ in order to find feasible allocations for (1.8) remains intractable in general. The suggested algorithm hence iteratively and greedily selects a single allocation $i_c^* \in \mathcal{I}^{c,(k)}, \forall c \in \mathcal{C}$. The main target is feasibility in (1.8), which is why in Lines 7 and 8 of Algorithm 2 the impact of choosing an allocation $i \in \mathcal{I}^{c,(k)}$ on the sum-power and the sum-rate constraints in (1.8b) and (1.8c) is evaluated. In Line 9 a column i_c^* is chosen which minimizes a weighted sum of sum-power and sum-rate constraint violations. Note that the scheme tries to prevent grave performance loss compared to $P^{*,\text{ts}}_{(\mathbf{R},\hat{\mathbf{P}})}$ by restricting this selection to a subset of columns $\breve{\mathcal{I}}^c \subseteq \mathcal{I}^{c,(k)}, c \in \mathcal{C}$, with objective values in (1.12) below a certain threshold. This restriction is successively relaxed in case no feasible allocation for the primal problem in (1.8) was found.

In the following we complete our framework by giving an overview of various heuristics for the discrete per-subcarrier problems in (4.3b) as used in Line 7 of

---

Algorithm 3 Joint Greedy Optimization (JOGO)

1: $[\mathbf{r}, \mathbf{p}(\mathbf{r}), f(\mathbf{p}(\mathbf{r}), \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda})] = \text{JOGO}(\mathbf{r}^0, U^{\text{opt}}, \boldsymbol{\lambda}, \boldsymbol{\nu})$

2: $\mathbf{r} = \mathbf{r}^0, \delta^* = 0, f^{\text{prev}} = f(\mathbf{p}(\mathbf{r}^0), \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda})$

3: **while** $\delta^* \leq 0$ **do**

4:   **for** $u = s_{U^{\text{opt}}+1}, \ldots, s_U$ **do**

5:     **if** $\exists \mathbf{p} \in \mathcal{Q} | r_u(\mathbf{p}) = r_u + \theta, r_i(\mathbf{p}) = r_i, \forall i \in \mathcal{U} \setminus \{u\}$,

6:     **then** $\delta_u = f(\mathbf{p}, \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda}) - f^{\text{prev}}$

7:     **else** $\delta_u = \infty$

8:   **end for**

9:   $u^* = \text{argmin}_{u=U^{\text{opt}}+1\ldots U} \delta_u, \ \delta^* = \delta_{u^*}$

10:   **if** $\delta^* \leq 0$ **then** $r_{u^*} = r_{u^*} + \theta, \ f^{\text{prev}} = f^{\text{prev}} + \delta_{u^*}$

11: **end while**

NDW-DSM in Algorithm 1. However, we hasten to add that the proposed decomposition framework is general and works, in principle, with any heuristic. As these subproblems can be studied independently we will drop the subcarrier index c throughout the rest of this chapter for ease of notation.

4.2 Heuristics for Discrete Rate Allocation

In Chapter 3 we studied the optimal solution of the subproblems in (4.3b) by branch-and-bound type of algorithms under a problem-specific variable-range reduction strategy. Using this scheme we observed that problem instances with around 16 users are optimally solvable in short time when crosstalk levels and/or target-rates are low. However, in general we found optimal solutions impractical for DSL networks with a large number of users, motivating suboptimal heuristics being used in our NDW-DSM framework. In the following we give an overview of the basic building blocks of a large class of more sophisticated meta-heuristics [181], and propose two more advanced, randomization-based heuristics.

4.2.1 Constructive Greedy Search Schemes

In Algorithm 3 we describe an iterative joint greedy optimization scheme (JOGO) for optimizing the per-subcarrier optimization problems (4.3b) over the bit-allocation of $U^{\text{greed}} = U - U^{\text{opt}}$ users, where we assume the allocations r_u of users $1 \leq u \leq U^{\text{opt}}$, are fixed. In each iteration the cost for loading another θ bits for any of the U^{greed} users is calculated and θ bits are allocated to the user minimizing this cost, cf. Lines 4–10 in Algorithm 3. Differently to JOGO, in the greedy heuristic SEGO the bit-allocation is performed sequentially over users, cf. Algorithm 4. Each user greedily minimizes the Lagrangian $f(\mathbf{p}, \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda})$ under fixed bit-allocations of already loaded users, cf. Line 5. Users loading bits at an earlier stage see in general

---

Algorithm 4 Sequential Greedy Optimization (SEGO)

1: $[\mathbf{r}, \mathbf{p}, f(\mathbf{p}, \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda})] = \mathrm{SEGO}(\mathbf{r}^0, U^{\mathrm{opt}}, \boldsymbol{\lambda}, \boldsymbol{\nu})$

2: $\mathbf{r} = \mathbf{r}^0$, $U^{\mathrm{greed}} = U - U^{\mathrm{opt}}$

3: Determine sequence $\mathbf{s} \in \mathbb{R}^{U^{\mathrm{greed}}}$ by ordering users $U^{\mathrm{opt}} + 1 \leq u \leq U$ in descending order of $(\check{w}_u + \lambda_u)/(\hat{w}_u + \nu_u)$

4: **for** $u = s_1, \ldots, s_{U^{\mathrm{greed}}}$ **do**

5: $\quad [r^u, \mathbf{p}] = \underset{\{r_u \in \mathcal{B}, \mathbf{p} \in \mathcal{Q} | \mathbf{r}(\mathbf{p}) \succeq \mathbf{r}\}}{\mathrm{argmin}} \{f(\mathbf{p}, \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda})\}$

6: **end for**
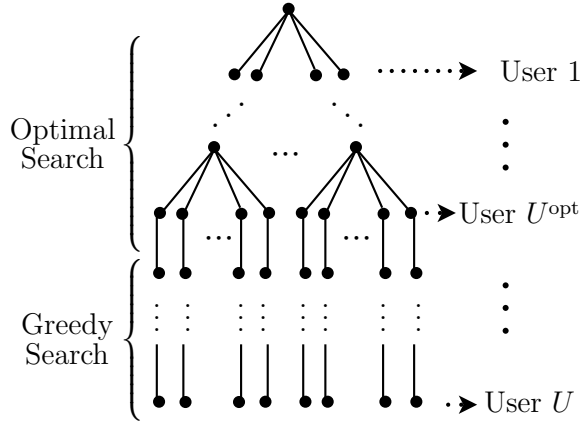
Figure 4.1: Schematic illustration of the bit-allocation search-tree in a mixed exhaustive and greedy line search.

less crosstalk and therefore encounter more possibilities for bit-loading than if their turn had been at a later stage. Therefore we heuristically reschedule the users based on their weights for rates and power using the relative metric $(\check{w}_u + \lambda_u)/(\hat{w}_u + \nu_u)$, cf. Line 3 in Algorithm 4. Comparing the two heuristics JOGO and SEGO from a complexity point of view and assuming a recursive computation of the matrix inverses for evaluating $\mathbf{p}(\mathbf{r})$ [103] we have that JOGO has a complexity per loaded bit-step θ of $O(U^3)$, while that of SEGO is only $O(U^2)$, cf. the cost-update in Lines 4–7 of Algorithm 3.

To round-off our description of greedy heuristics we note that they can also be applied jointly with optimal branch-and-bound schemes as in [163] and Chapter 3, for instance to take advantage of the presence of a few dominant disturbers. More precisely, an exhaustive search can be represented by a search tree where level $u \in \mathcal{U}$ of the tree relates to the bit-loading decision of the u'th user and the leaves of the tree correspond to the discrete power-allocations $\mathcal{Q}$, cf. Section 3.1. We can make a mixed exhaustive and greedy search (MEGS) by only performing an exhaustive search for the first U^{opt} users, while for each tested allocation $r_u, 1 \leq u \leq U^{\mathrm{opt}}$, a heuristic algorithm is used to allocate bits to the remaining $U^{\mathrm{greed}} = U - U^{\mathrm{opt}}$ users,

---

Algorithm 5 Local Search (LS) based Bit-Loading

1: Initialize $\mathbf{r}$
2: **repeat**
3:    Update $\mathbf{r}$ by
4:    **a)** the first-found $\tilde{\mathbf{r}} \in \mathcal{N}(\mathbf{r})$ with $f(\tilde{\mathbf{r}}) < f(\mathbf{r})$, or
5:    **b)** any $\tilde{\mathbf{r}} \in \mathcal{N}(\mathbf{r})$ with $f(\tilde{\mathbf{r}}) < f(\mathbf{r})$, $f(\tilde{\mathbf{r}}) \leq f(\bar{\mathbf{r}})$, $\forall \bar{\mathbf{r}} \in \mathcal{N}(\mathbf{r})$
6: **until** Convergence

---

cf. Figure 4.1. While there are various options on the design of the optimal and the heuristic search part, in our simulations we use a depth-first branch-and-bound scheme from Chapter 3 and SEGO in Algorithm 4, respectively. In Section 4.3.1 we will study a near-far scenario in which the proposed decomposition of the search tree in MEGS leads to a near-optimal solution at a reduced complexity compared to optimal allocation schemes.

4.2.2 Local search

An essential part of many meta-heuristics is the local search (LS), where in [181] we found that a simple scheme presented next is able to substantially improve the average performance of the presented greedy schemes. In local search schemes one iteratively moves from an allocation $\mathbf{r}^{(\tilde{k})}$ in iteration $\tilde{k}$ of the search to an improving allocation $\mathbf{r}^{(\tilde{k}+1)} \in \mathcal{N}(\mathbf{r}^{(\tilde{k})}) \subseteq \mathcal{B}$ where $f(\mathbf{r}^{(\tilde{k}+1)}) < f(\mathbf{r}^{(\tilde{k})})$. The algorithm terminates when no such improving step is possible, i.e., when a local optimum $\mathbf{r}$ with $f(\tilde{\mathbf{r}}) \geq f(\mathbf{r}), \forall \tilde{\mathbf{r}} \in \mathcal{N}(\mathbf{r})$ has been reached, cf. Algorithm 5. The set $\mathcal{N}(\mathbf{r})$ is called the neighborhood of $\mathbf{r}$, where we found [181] that a simple but effective choice is the set

$$\mathcal{N}^{(2)}(\mathbf{r}) = \mathcal{N}^{(1)}(\mathbf{r}) \cup \bar{\mathcal{N}}^{(2)}(\mathbf{r}), \tag{4.8a}$$

$$\mathcal{N}^{(1)}(\mathbf{r}) = \{\tilde{\mathbf{r}} \in \prod_{u \in \mathcal{U}} \mathcal{B} \mid \tilde{r}_u = r_u \pm \theta, \tilde{r}_i = r_i, \forall i \in \mathcal{U} \setminus \{u\}, u \in \mathcal{U}\}, \tag{4.8b}$$

$$\bar{\mathcal{N}}^{(2)}(\mathbf{r}) = \{\tilde{\mathbf{r}} \in \prod_{u \in \mathcal{U}} \mathcal{B} \mid \tilde{r}_u = r_u \pm \theta, \tilde{r}_{\bar{u}} = r_{\bar{u}} \pm \theta, \tag{4.8c}$$

$$\tilde{r}_i = r_i, \forall i \in \mathcal{U} \setminus \{u, \bar{u}\}, u \neq \bar{u}, u, \bar{u} \in \mathcal{U}\}, \tag{4.8d}$$

which contains all allocations that can be reached by perturbing at most two elements of $\mathbf{r}$. Note that the complexity of a local search depends on the initialization point $\mathbf{r}^{(0)}$ and is intuitively lower when the search is initialized "close" to a local optimum. This observation leads to a "warm-start" local search scheme described in Section 4.2.7. The following result characterizes the asymptotic size of the proposed neighborhood in U and the complexity of local search, respectively.

Theorem 9. *Assuming $\frac{H^{ui}}{H^{uu}} \geq \alpha > 0, \forall u, i \in \mathcal{U}$, the numbers of neighboring points*

$|\mathcal{N}^{(1)}(\mathbf{r})|$ *and* $|\mathcal{N}^{(2)}(\mathbf{r})|$ *to a point* $\mathbf{r}$ *with* $\mathbf{p}(\mathbf{r}) \in \mathcal{Q}$ *grow as* $O(U)$, *respectively.*

Proof. The first part on $|\mathcal{N}^{(1)}(\mathbf{r})|$ follows trivially from the definition in (4.8b), while the second part on $|\mathcal{N}^{(2)}(\mathbf{r})|$ follows from the proof of Theorem 3 as follows. The number of users $u \in \mathcal{U}$ with non-zero bit-allocation $r_u > 0$ in a feasible allocation $\mathbf{r}$ with $\mathbf{p}(\mathbf{r}) \in \mathcal{Q}$ is by Theorem 3 bounded by a constant $\hat{U}$ under the assumptions of the theorem. All those users $u \in \mathcal{U}$ with $r_u = 0$ can only increase their rates. Therefore, the size $|\bar{\mathcal{N}}^{(2)}(\mathbf{r})|$ of the set in (4.8c) comprising all allocations generated by changing *exactly* two elements of $\mathbf{r}$ is bounded by $O(\hat{U}U)$. Altogether the size $|\mathcal{N}^{(2)}(\mathbf{r})| = |\mathcal{N}^{(1)}(\mathbf{r})| + |\bar{\mathcal{N}}^{(2)}(\mathbf{r})|$ of the set in (4.8a) grows linearly in U. □

Corollary 5 (of Thm. 3 and Thm. 9). *The local search for the problem in (4.3b) has polynomial complexity in U under the assumptions and the neighborhood set of Theorem 9.*

Proof. This follows from Theorems 3 and 9 and by the polynomial complexity of evaluating $\mathbf{p}(\mathbf{r})$ by solving the linear system in (1.5). □

The assumptions in Theorem 9 are satisfied in the following. More generally they hold in all randomized schemes in Sections 4.2.4 and 4.2.5 and in [181] as these only evaluate the neighborhood around rates $\mathbf{r}$ where the power-allocation $\mathbf{p}(\mathbf{r})$ is in the set of feasible PSD's $\mathcal{Q}$.

4.2.3 Analysis of the Constructive Greedy Heuristics

We analyze our two greedy heuristics by simulation on a large set of deterministically generated problem instances: We choose a subset of subcarriers $\tilde{\mathcal{C}} = 1, 51, \ldots, 1601$, fixed Lagrange multipliers $\nu_u = 1/U, \lambda_u = 1, u \in \mathcal{U}$, and construct our network scenarios using a set of specified line lengths $\{200, 400, 600, 800\}$ m, considering all U-combinations with repetitions, cf. Section 3.3.2. The simulation parameters were selected according to the VDSL standard in [50], with $\Gamma = 12.8$ dB, $\hat{\theta} = 15$, $\theta = 1$, upstream band plan 997-M1x-M, and noise N_u^c resulting from a summation of VDSL noise A [50] added to a flat background noise at -140 dBm/Hz. The maximum and average suboptimality of JOGO over subcarriers $\tilde{\mathcal{C}}$ and scenarios is 39.3 % and 3.7 %, respectively. While the suboptimality values of JOGO were found to be zero for all collocated scenarios, the highest values appear in classical near-far type of scenarios, cf. [180] for details. For example, the largest suboptimality of 13.1 % appears for 4 users located at 800 m and 1 user located at 200 m and 400 m each. To analyze this example even further we pick a specific subcarrier at approximately 9.26 MHz. The greedy base heuristic JOGO assigns the bits $\mathbf{r} = [9, 4, 0, 0, 0, 0]$, i.e., only the shortest lines are transmitting. Compared to an optimal allocation $\mathbf{r}^* = [6, 4, 3, 2, 2, 2]$ JOGO has a suboptimality of more than 31 %. Note that in this example SEGO gives the result $\mathbf{r} = [16, 0, 0, 0, 0, 0]$ when the user sequence starts with the shortest line, corresponding to a suboptimality of more than 15 %. Also, having the two

greedy algorithms followed by the local search in Algorithm 5 does not improve the solutions beyond 15 % of suboptimality. We noted that any user sequence in SEGO which starts with one of the longer lines would have lead to an optimal result after a following local search initialized at the solution of SEGO. This motivates the extension of SEGO in Section 4.2.5 where the user-sequence is optimized together with the bit-allocation. In Section 4.2.4 we present a modification of JOGO which uses randomized greedy decisions and local search. These two heuristics presented next were seen to outperform those in Section 4.2.1 as well as various other meta-heuristics [180].

4.2.4 A Greedy Randomized Adaptive Search Procedure (Grasp)

As the name indicates, Grasp [139] is a randomized extension of greedy search schemes, cf. Algorithm 13. More precisely, Lines 3–12 mimic the joint greedy bit-loading of JOGO with the important difference that in Line 11 a randomized, potentially non-greedy decision is taken. The parameter β allows for a trade-off between standard greedy schemes ($\beta = 0$) and purely random schemes ($\beta = 1$) [139]. However, just as in JOGO the greedy bit-loading process is considered complete when even a purely greedy decision would increase the solution objective, i.e., $\delta^{\min} > 0$, cf. Line 3. The free parameters in the presented basic implementation of Grasp are the local search strategy and neighborhood definition, cf. Line 14, the number of restarts K (where an alternative stopping criterion will be used in our simulations), and the randomization parameters $\beta_i, i \in \{1, \ldots, M\}$, where in each iteration of Grasp a single parameter β is chosen depending on the average objective value experienced under all parameters $\boldsymbol{\beta}$, cf. Lines 4 and 16 of Algorithm 13, respectively.

4.2.5 Randomized SEGO (rSEGO)

Next we present an improvement over SEGO in Section 4.2.1 in the sense that the sequence with which the users are loaded is adaptive, cf. Algorithm 14. It combines elements of SEGO, Grasp, as well as ant colony system heuristics [46], cf. [180] for an ant colony system based bit-loading scheme. More precisely, in Lines 9–14 of Algorithm 14 the currently active user u decides which user should allocate his bits next, where each possible decision $i \in \tilde{\mathcal{U}}$ is assigned a sequence decision value $\tilde{\tau}_u(i)$. The decision on the bit-rate r_u is made greedily in Line 7 similarly as in SEGO, where similarly as in Grasp the decision is randomized based on the objective values of all possible rate decisions. Finally, after $\bar{K}$ solutions have been constructed in this way, defining $\hat{f} = \sum_{u \in \mathcal{U}} (\hat{w}_u + \nu_u) \hat{p}_u$ (the largest possible objective value) and using any $\rho \in [0, 1]$, a global update of the decision values $\tilde{\tau}_u(j)$ is done in Line 19 as

$$\tilde{\tau}_u(j) = (1 - \rho) \cdot \tilde{\tau}_u(j) + \rho \cdot \left(\hat{f} - f(\mathbf{r}^{(k^*)}) \right). \tag{4.9}$$

Table 4.1: Improvements by randomized heuristics compared to JOGO.

Alg.	Mean objective improvement [%] (per subcarrier)	[Min.,Max.] improvement [%] of sum-objective over subcarriers	Mean op.-count $[\times 10^3]$ per subcarrier[7]
JOGO	0	$[0,0]$	1.79 ± 0.06
JOGO+LS	8.03 ± 0.78	$[0.91,32.13]$	3.92 ± 0.07
rSEGO$^{(10*)}$	8.90 ± 0.00	$[0, 29.024]$	10^*
rSEGO$^{(20*)}$	9.76 ± 0.00	$[1.36, 21.89]$	20^*
Grasp$^{(10*)}$	9.05 ± 0.00	$[0, 29.02]$	10^*
Grasp$^{(20*)}$	9.86 ± 0.00	$[1.48, 21.89]$	20^*

4.2.6 Simulations on Single-Carrier Heuristics

We will use the same simulation parameters and network generation method as in Section 4.2.3. Furthermore, we make the practical assumption that there is a restriction in simulation time for solving the subproblems in (4.3b). In order to make our results reproducible for future research, as in Section 4.3 we will use the number of power evaluations $\mathbf{p}(\mathbf{r})$ by solving the linear system in (1.5) as the stopping criterion for Grasp and rSEGO. This is further motivated by the almost identical simulation times of these heuristics as a function of the number of power evaluations. We will initialize the incumbent solution (but not the initial starting-point $\mathbf{r} = \mathbf{0}$) of all methods by the result of the base heuristic JOGO. As this heuristic is guaranteed to give a solution with negative objective value, we have that also all other heuristics will produce a solution $\mathbf{r}^*$ with negative objective value $f(\mathbf{r}^*)$. The choice of parameters[5] is based on Monte-Carlo simulations in six-user networks where optimal solutions could still be obtained by the optimal methods presented in Chapter 3. Comparing the two greedy randomized schemes rSEGO and Grasp, both, their performance metrics as well as the scenarios where their performance was suboptimal were found to be similar [181]. Grasp is based on JOGO and we found that it takes a parameter $\beta = 1$ to remedy the shortcomings of JOGO in the example made in Section 4.2.3.[6] We chose to have two parameters $\boldsymbol{\beta} \in \mathbb{R}^2$ among which we select, cf. Line 4 in Algorithm 13.

Due to the large number of 30-user scenarios for the four selected line-lengths we uniformly sample 200(100) scenarios (not cable lengths) for the deterministic (ran-

---

[5]For all randomized heuristics we chose a first-improving local search with the two-step neighborhood in (4.8a). Parameters for rSEGO were $\beta = 0.75$, $\bar{K} = 5$, $\rho = 0.99$, and $q_0 = 0.2$, while for Grasp we use $\boldsymbol{\beta} = [0.75, 1]$.

[6]Lines are not identical in their transmission parameters in practice. Still, the example suggests β to be set to a large value in near-far scenarios.

[7]The numbers marked by * reflect the set complexity budget.

domized) algorithms, make 50 repetitions for randomized algorithms, and present mean values with 95 % confidence intervals according to a t-test. In order to reduce the variance of our results we ran all heuristics on identical, sampled scenarios. Furthermore, as we lack an optimal solution we use the greedy algorithm JOGO as our reference and show the improvements in objective value. Table 4.1 reports our average simulation results, where the values marked by a star reflect the applied stopping criterion (complexity budget) and exclude the incumbent initialization using JOGO. We find that using a combination of JOGO and local search (with a negligible difference whether we initialized LS at the solution of JOGO or at $\mathbf{r} = \mathbf{0}$, using JOGO solely as an initial incumbent solution) already gives improvements compared to the objective value under JOGO of around 8 %, while the average improvements of our randomized algorithms are notably almost 10 % under the limit of $2 \cdot 10^3$ matrix inversions.

4.2.7 Warm-Start Local Search (WS-LS)

Despite the presented low-complexity heuristics, solving the single-carrier problem in (4.3b) on all subcarriers $c \in \mathcal{C}$ leads to a large total complexity. We therefore propose a novel scheme where only a subset of all subcarriers $\tilde{\mathcal{C}} \subseteq \mathcal{C}$ is solved using one of the single-carrier heuristics. The approach WS-LS takes to obtain a reasonable starting point $\bar{\mathbf{r}}$ for local search on all other subcarriers $c \in \mathcal{C} \setminus \tilde{\mathcal{C}}$ is to use the heuristic solutions $\mathbf{r}^{\hat{c}}$ and $\mathbf{r}^{\check{c}}$ on the "neighboring" subcarriers $\hat{c}, \check{c} \in \tilde{\mathcal{C}}$. The conservative strategy we suggest to use is $\bar{r}_u = \min\{r_u^{\hat{c}}, r_u^{\check{c}}\}, \forall u \in \mathcal{U}$. Note that WS-LS is different to the idea of subcarrier grouping in wireless networks or the complexity reduction technique in [66] as we do not approximate or interpolate solutions on different subcarriers. The following results are based on a maximum of 10^3 power evaluations per subcarrier, a selection of subcarriers $\tilde{\mathcal{C}}$ in regular intervals, and other settings as specified in Section 4.2.6. As a base-line we added a randomized local search (rLS) scheme where the LS algorithm is reinitialized at random starting points uniformly drawn from the set of allocations feasible w.r.t. single-user bit-cap and power mask constraints. Figure 4.2 depicts the average improvements in sum-objective by WS-LS and its complexity in all 6-user scenarios as described in Section 4.2.6 compared to using JOGO on all subcarriers. These results show that the performance gain by increasing $|\tilde{\mathcal{C}}|$ actually flattens out, with this effect happening for lower values of $|\tilde{\mathcal{C}}|$ using Grasp or rSEGO compared to rLS, cf. Figure 4.2(a). The initial drop in complexity in Figure 4.2(b) is explicable by the dependence of the local search complexity on its starting point. As an example, we find that using WS-LS with $|\tilde{\mathcal{C}}| = 21$ we can achieve a complexity reduction of more than 90 % compared to using rSEGO with given complexity limitation.
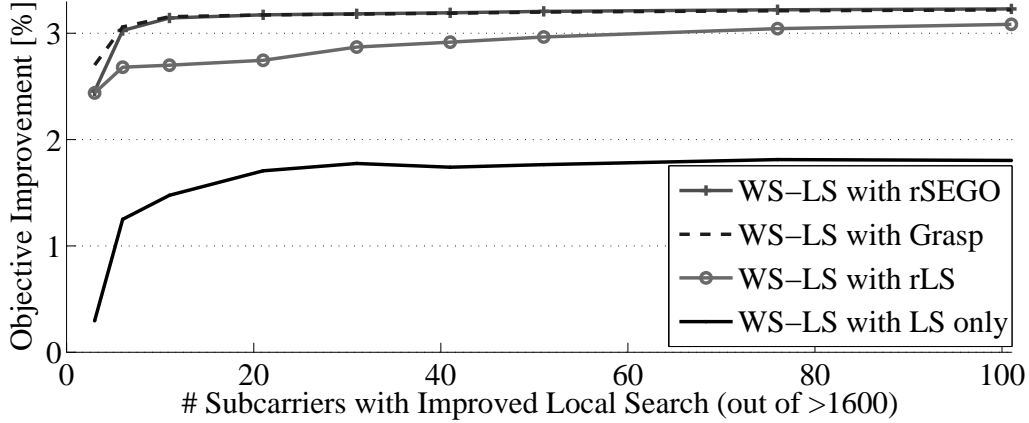
(a)
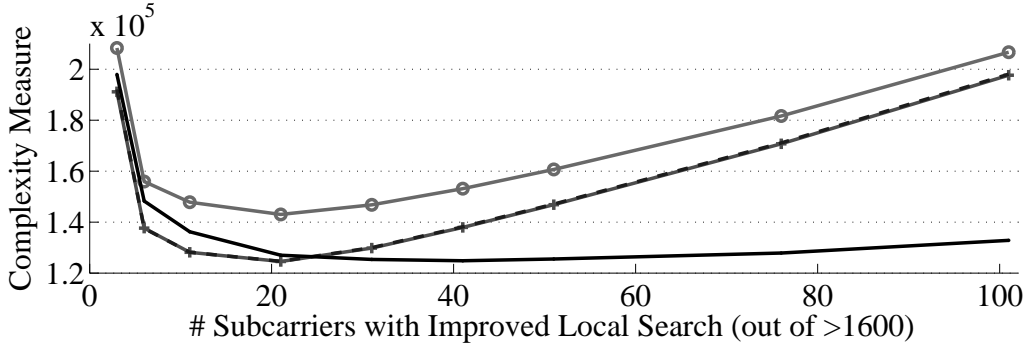
(b)

Figure 4.2: Trade-off for heuristic WS-LS between average objective improvement compared to JOGO (a) and sum-complexity over all C subcarriers (b).

4.3 Simulation Results on Low-Complexity DSM

In Section 4.3.1 we will demonstrate the marginal loss incurred by CHET, the heuristic for recovering primal feasible solutions from the solution of NDW-DSM. At the same time we show an application scenario for the MEGS heuristic. The simulation results in Section 4.3.2 give an example of a large network with 50 lines where our proposed DSM framework gives substantial performance gains compared to previous DSM schemes. In order to demonstrate that our algorithm performs very well not only in special cases, we look at the *average* sum-rate performance achieved under our framework compared to previous algorithms in a set of 10^3 distributed DSL scenarios in Section 4.3.3. Throughout this section we assume a DSL system with
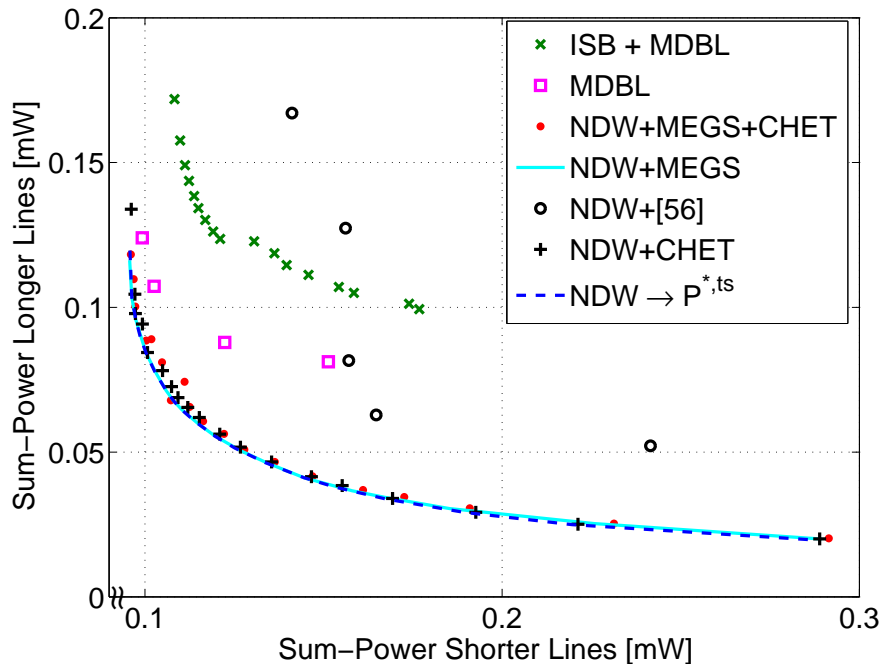
Figure 4.3: Sum-power regions in a 4-user VDSL upstream near-far scenario.

simulation parameters chosen as in Section 4.2.3. We will compare our algorithm to single-user bit-loading [27] under worst-case crosstalk noise $\sum_{i \in \mathcal{U} \setminus u} H_c^{ui} \hat{p}_c^i, c \in \mathcal{C}$, for user $u \in \mathcal{U}$ ("Mask-Based" SSM), to an iterative spectrum balancing (ISB) algorithm[8], and the multi-user DBL (MDBL) algorithm in [103].

4.3.1 A Near-Far DSL Example - Applying MEGS and CHET

In Figure 4.3 we present the power-regions of various DSM algorithms in a 4-user VDSL upstream near-far scenario.[9] In this scenario we have that the crosstalk between the longer two lines is low compared to the shorter two lines and we therefore

---

[8]In ISB we perform the Lagrange multiplier update sequentially over users while fixing PSDs as in [31] but differently use a bisection search for this purpose and perform the line-search over the bit-rates. As ISB does not result in a discrete bit-allocation we floor the final bits and run MDBL [103] from this initialization, cf. [196] for a similar approach in extending continuous rate-maximizing DSM schemes. The convergence criterion is a maximum number of 50 user sweeps not improving the Lagrangian or a total of 200 iterations, while for the number of iterations in NDW-DSM using heuristics we set $30 \le k \le 60$ and set $\bar{\bar{\delta}}$ to $0.1\,\text{ppm}$, cf. Line 2 in Algorithm 1. The used LP solver for the problem in (4.1) is the primal-dual interior point solver in [120].

[9]Parameters for Figure 4.3 were $R_1 = R_2 = 42\,\text{Mbps}$, $R_3 = R_4 = 3\,\text{Mbps}$, $\hat{w}_1 = \hat{w}_2 = 0.5 * [0.05, 0.1, \dots, 0.95]$, $\hat{w}_3 = \hat{w}_4 = 0.5 * [0.95, 0.9, \dots, 0.05]$, $\breve{w}_u = 0, \forall u \in \mathcal{U}$, shorter and longer two lines have a length of $300\,\text{m}$ and $1000\,\text{m}$, respectively, and are collocated at the CO side. Results for the upper three curves show non-dominated points only.

expect the MEGS heuristic presented in Section 4.2.1 to be useful. As Figure 4.3 shows, in this near-far scenario the power-region obtained by the MEGS heuristic is close to the lower-bound given by the optimum[10] $P^{*,\text{ts}}$ of the time-sharing problem in (2.3), and larger compared to the other two popular heuristics, even when we employ CHET on top. For example, the lowest sum-power obtained by the combination of ISB with MDBL is more than 30 % above that obtained by the combination of NDW-DSM with MEGS and CHET, while applying MDBL alone implies an increase of more than 19 %. This also shows that initializing bit-loading schemes using other (e.g., continuous) DSM schemes does not necessarily improve their performance.

The share of subcarriers where time-sharing is applied at the solutions of the time-sharing problem in (2.3) is in the range between 0 % (meaning that a single solution was found and the relaxation gap between the problems in (1.8) and (2.3) is hence zero) and 60 %. These numbers depend however on the used LP solver. While we apply the interior-point solver in [120], a simplex solver [14] would result in a number of time-shared subcarriers as given in Theorem 3.

Figure 4.3 also exemplifies the performance loss when the heuristic in [53] is used jointly with NDW-DSM, cf. Section 4.1.4 for an intuitive explanation.[11] Differently, the CHET heuristic for recovering from time-sharing has, for the presented results of this section and in combination with NDW-DSM, a guaranteed suboptimality in the original problem (1.8) of less than approximately 2 %. This is inferred by using the lower-bound on the optimal objective of (1.8) obtained by optimally solving the time-sharing relaxation in (2.3), cf. the lowest curve in Figure 4.3. This shows the validity of our approach to optimize (1.8) by (e.g., approximately) solving the time-sharing problem in (2.3) and reconstructing solutions for (1.8) by the CHET heuristic.

4.3.2 Mixed Near-Far DSL Example with 50 Users

Here we consider a 50-user upstream mixed VDSL scenario as shown in Figure 4.5 with $L_{CO-Cab} = 500\,\text{m}$, $L_{Cab-u} = 300\,\text{m}$, $\forall u \in \mathcal{U}$, where half of the users are deployed from the central office and the other half from the cabinet. Figure 4.4 shows the results of various DSM schemes under rate-maximization ($\hat{\mathbf{w}} = \mathbf{0}, \check{\mathbf{w}} = \mathbf{1}$, cf. (1.9)) and sum-power minimization ($\hat{\mathbf{w}} = \mathbf{1}, \check{\mathbf{w}} = \mathbf{0}$, $R_u = 6\,\text{Mbps}$, $\forall u \in \mathcal{U}$), respectively. We apply NDW-DSM using the warm-start local search heuristic and rSEGO. The runtime of NDW-DSM in this scenario was around 6 hours, while that of ISB and

---

[10]The time-sharing problem in (2.3) is solved optimally by Algorithm 1, solving the subproblems in (4.3b) as stated in Lines 5 and 6 in Algorithm 1, and omitting Line 10 which only concerns the original problem in (1.8).

[11]We adapted the mixing algorithm in [53] to the studied sum-power minimization problem by using the (normalized [53]) Euclidean distance to the optimal time-shared solution found by NDW-DSM as the selection metric.
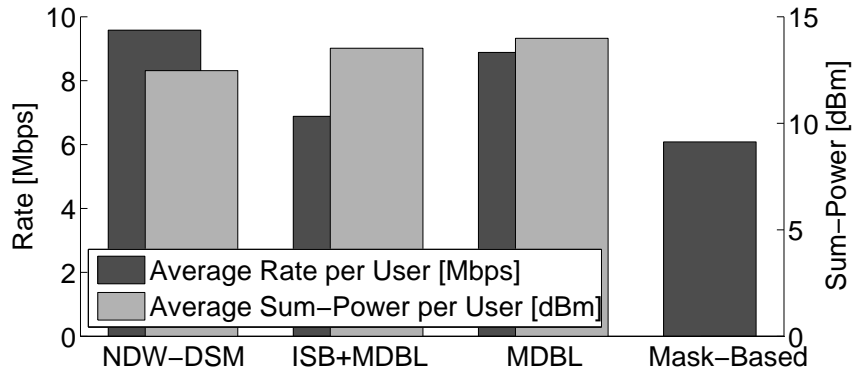
Figure 4.4: Average per-user rate and sum-power in a 50-user VDSL scenario.[12]
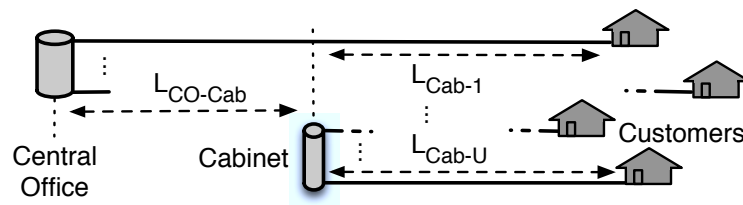
Figure 4.5: Schematic of a mixed near-far DSL deployment.

MDBL was in the order of 1.3 and 0.4 hours, respectively.[13] However, in terms of sum-rate NDW-DSM shows a gain of 57.5 %, 39.1 % and 7.8 % in comparison to "Mask-Based" SSM, ISB and MDBL, respectively. Regarding solely the short lines NDW-DSM improves their sum-rate by 88.4 % compared to MDBL, demonstrating the disadvantage of long lines under the greedy DSM scheme MDBL. In terms of transmit sum-power NDW-DSM saves 21.6 % and 29.5 % compared to ISB and MDBL, respectively. In conclusion, this example demonstrates the possible gains of our combinatorial search compared to user-iterative and greedy DSM schemes under discrete rates.

---

[13]All methods are coded in Matlab with the exception of the local search in NDW-DSM and the line-search in ISB which are written in C. The platform is an 8-core Intel system at 3.33 GHz with 12 GB RAM. The subproblem solutions in NDW-DSM for subcarrier groups as suggested by the warm-start method in Section 4.2.7 were parallelized in Matlab over 4 processes. We note that the exact runtimes may vary depending on the exact channel model. For example, in experiments using the measurements in [153] and assuming the lines are distributed in three binders with an inter-binder attenuation of 7.6 dB as reported in [115] we observed roughly 50 % higher simulation times for rate maximization. We attribute this behavior to the fairly low crosstalk couplings and therefore larger number of feasible allocations compared to the worst-case crosstalk model, cf. the discussion on complexity in Section 2.1.2.

[13]Sum-power results under the mask-based scheme were omitted as the target-rates could not be reached.
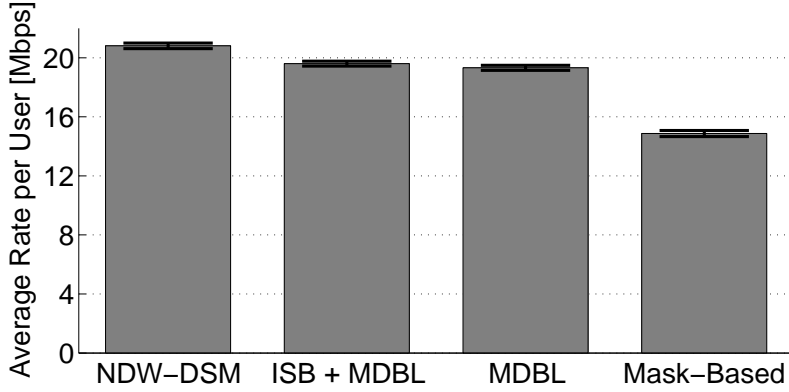
Figure 4.6: Average rates by various DSM schemes in 1000 DSL scenarios with 99 % confidence intervals.

4.3.3 Average Performance for Mixed Near-Far DSL Deployments

Differently to the specific DSL scenario investigated in the previous section, we will now analyze the performance of our DSM framework in a large set of mixed near-far VDSL deployments as shown in Figure 4.5 with 25 lines out of which 10 lines connect to the cabinet. We generated 10^3 downstream scenarios by uniformly sampling the lengths $L_{CO-Cab} \in [100, 1400]$ and $L_{Cab-u} \in [50, 500], \forall u \in \mathcal{U}$. Note that this random topology selection results in a diverse set of generated crosstalk coupling scenarios. The average runtime complexity of NDW-DSM per network scenario was approximately one hour and therefore in the order of the runtime complexity of ISB, but significantly higher than that of MDBL which on average only required in the order of 5 minutes. However, we note that reductions in runtime of NDW-DSM as well as ISB may be achieved by further parallelization of the per-subcarrier subproblem solutions.

Figure 4.6 shows the average per-user rates achieved by different DSM schemes together with 99 % confidence intervals according to a t-test. The average gain by NDW-DSM is $+1.21 \pm 0.03$ Mbps (6.2 %), $+1.49 \pm 0.05$ Mbps (7.7 %), and $+5.94 \pm 0.06$ Mbps (40.0 %) compared to ISB, MDBL and Mask-Based SSM, respectively. Regarding once more only the central-office deployed lines we have an average gain by NDW-DSM compared to MDBL of $+1.56 \pm 0.05$ Mbps (11.1 %), confirming the behavior of MDBL towards longer lines as observed in Section 4.3.2.

Chapter 5 Energy-Efficient Maximization of Service Coverage

In this chapter we propose and analyze a novel DSM problem, namely the maximization of the service coverage. Coverage will be defined as the number of DSL lines which can be granted an operator-specified high-bandwidth service. The network operators think in terms of the number of customers they can offer a particular service, which is better reflected by this objective than by previous DSM formulations. Furthermore, we hypothesize that controlling the users' powers to optimize coverage also leads to a sum-power reduction. This intuition is based on the fact that the over-performing users' rates are reduced to the service rate and the rates of under-performing users which can not achieve the service rate are reduced to a pre-defined minimum rate. Current objectives for DSM in the literature range from sum-rate [197, 199] and min-rate (or reach) maximization [199], to sum-power minimization [164, 186] and general utility (e.g., fairness [93, 145, 165] or hardware power consumption [66]) optimization, cf. also Sections 1.2.4, 2.3.1, and 4.1.3 for various examples. As we deem our original problem formulation including binary service variables intractable for large DSL networks we initially apply a convex relaxation of these integer variables, from which we recover after rate-allocation and power-allocation in a variable fixing phase. In Section 5.1 we compare two potential such relaxations, namely a linear continuous relaxation and a semidefinite program relaxation [102, 169]. We present simulation results in a large number of small-size network scenarios where we can still compare to the *global* optimum. Irrespective of the applied relaxation our heuristic approach for coverage maximization (CM) is based on a modification of the column generation framework of Chapter 4, where we exploit the method's memory of numerous per-subcarrier solutions in order to test various integer fixings of the relaxed binary variables at comparably low complexity. This strategy also allows us to pursue energy-efficiency by reducing the users' rates according to the found service coverage, namely by having the service coverage optimization followed by a sum-power minimizing DSM optimization. In other words, we swap the optimization objective after having settled on a certain set of target-rates. In Section 5.2.2 we show simulation results in 1000 25-user near-far DSL scenarios which exemplify the substantial gain in service coverage by our heuristic compared to various sum-rate maximizing DSM algorithms, and also support our intuition on

the effect of CM on the sum-power consumption.

5.1 Problem, Relaxations, and a Heuristic

5.1.1 An Integer Problem Formulation

Our optimization problem is to guarantee a basic target-rate for all users while obeying all system constraints and maximizing the number of users which can be upgraded to a high-bandwidth service. Denoting the decision variables for service upgrades by $b_u \in \{0, 1\}, u \in \mathcal{U}$, we formulate this coverage maximization (CM) problem for DSL as

$$P^{cov}_{(\mathbf{R},\hat{\mathbf{P}})} = \max_{\substack{\mathbf{p}^c \in \mathcal{Q}^c, c \in \mathcal{C}, \\ b_u \in \{0,1\}, u \in \mathcal{U}}} \sum_{u \in \mathcal{U}} b_u \tag{5.1a}$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} r_u^c(\mathbf{p}^c) \geq R_u + b_u(\hat{R} - R_u), \quad \forall u \in \mathcal{U}, \tag{5.1b}$$

$$\sum_{c \in \mathcal{C}} p_u^c \leq \hat{P}_u, \quad \forall u \in \mathcal{U}, \tag{5.1c}$$

where $\hat{R} \in \mathbb{R}_+$ is the target-rate for the service upgrade, and where $\mathbf{R} \in \mathbb{R}_+^U$ and $\hat{\mathbf{P}} \in \mathbb{R}_+^U$ are the users' minimum target-rates and maximum sum-powers, respectively. As the variables $b_u, u \in \mathcal{U}$, are binary the objective in (5.1a) reflects the number of high-bandwidth users, i.e., the service coverage. We proceed by applying a relaxation of the binary variables similarly as in the time-sharing relaxation of Section 2.1.2, leading to the linear relaxation ("lr")

$$P^{cov,lr}_{(\mathbf{R},\hat{\mathbf{P}})} = \max_{\substack{\mathbf{p}^c \in \mathcal{Q}^c, c \in \mathcal{C}, \\ b_u \in [0,1], u \in \mathcal{U}}} \sum_{u \in \mathcal{U}} b_u \tag{5.2a}$$

$$\text{subject to} \quad \text{Constraints in (5.1b) and (5.1c)} \tag{5.2b}$$

Continuous linear (i.e., LP) relaxations are the standard approach in mixed integer programming, typically applied in conjunction with cut generation procedures which tighten the relaxation by "cutting off" non-integral solutions from the feasible constraint set [124]. We will investigate an alternative relaxation in Section 5.1.2 which is based on a lifting of the binary variables into a higher dimensional space. Also, we will present a method to return to a feasible binary solution in Section 5.1.3.

Assuming $\hat{R}$ is a multiple of the bit-step size θ, due to $b_u \in \mathbb{R}$ in this linear relaxation we have that the rate constraints of the linear relaxation in (5.2) hold with equality at optimum. Hence, by reformulation of these constraints we find that the linearly relaxed problem in (5.2) can be interpreted as a weighted sum-rate maximization problem with weights $\breve{w}_u = (\hat{R} - R_u)^{-1}$, $u \in \mathcal{U}$, and maximum sum-

rate constraints additional to the minimal target-rate constraints. In other words, the priority of a user $u \in \mathcal{U}$ is larger the smaller the difference between the service rate $\hat{R}$ and the user's target-rate R_u is. We will study the impact of setting such minimal and maximal rate constraints on the service coverage by simulations in Section 5.2.2.

Based on the same motivation as in Section 2.1.2 we further relax the problem in (5.2) by a time-sharing relaxation of the rate and power-allocations, leading to a linear master problem similar to that in (2.3), given as

$$P^{cov,lr,ts}_{(\mathbf{R},\hat{\mathbf{P}})} = \underset{\xi^c_i \geq 0, i \in \mathcal{I}^c, c \in \mathcal{C}, b_u \in [0,1], u \in \mathcal{U}}{\text{maximize}} \sum_{u \in \mathcal{U}} b_u \tag{5.3a}$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}^c} r^c_u(\mathbf{p}^{c,i}) \xi^c_i \succeq R_u + b_u(\hat{R} - R_u), \forall u \in \mathcal{U}, \tag{5.3b}$$

$$\text{Constraints in (2.3c) and (2.3d).} \tag{5.3c}$$

By the same motivation as in Section 2.1.2 we suggest the application of the column generation process described in Section 4.1.1, upon termination of which the relaxed variables $b_u, u \in \mathcal{U}$, will provide heuristic priority information on how to round these variables to integer values. Before developing this idea in more detail we next regard an alternative to the applied linear relaxation.

5.1.2 Semidefinite Relaxation of the Service Coverage Problem

The binary constraint $b_u \in \{0, 1\}, u \in \mathcal{U}$, implies the relation $b^2_u = b_u, u \in \mathcal{U}$, and we can therefore equivalently rewrite the original problem in (5.1) as

$$\underset{\substack{\mathbf{p}^c \in \mathcal{Q}^c, c \in \mathcal{C}, \\ b_u, u \in \mathcal{U}}}{\text{maximize}} \quad \sum_{u \in \mathcal{U}} b_u \tag{5.4a}$$

$$\text{subject to} \quad b^2_u = b_u, \quad \forall u \in \mathcal{U}, \tag{5.4b}$$

$$\sum_{c \in \mathcal{C}} r^c_u(\mathbf{p}^c) \geq R_u + b^2_u(\hat{R} - R_u), \quad \forall u \in \mathcal{U}, \tag{5.4c}$$

$$\text{Constraint (5.1c).} \tag{5.4d}$$

The semidefinite programming (SDP) relaxation [102] of this problem can be seen to be its Lagrangian bi-dual problem, as stated in Theorem 10, cf. Appendix A.9 for the common ad-hoc derivation [102] of the SDP relaxation with and without time-sharing relaxation, and Theorem 11 for some of its properties.

Theorem 10. *The bi-dual of the problem in (5.4) is the time-sharing/semidefinite*

relaxation given as

$$P_{(\mathbf{R},\hat{\mathbf{P}})}^{cov,sdp,ts} = \underset{\substack{\xi_i^c \geq 0, i \in \mathcal{I}^c, c \in \mathcal{C}, \\ \breve{\boldsymbol{B}} \in \mathbb{R}^{(U+1)\times(U+1)}}}{\text{maximize}} \sum_{u \in \mathcal{U}} \breve{B}_{uu} \tag{5.5a}$$

$$\text{subject to} \quad \breve{B}_{uu} = \breve{B}_{(U+1)u}, \forall u \in \mathcal{U}, \;\; \breve{B}_{(U+1)(U+1)} = 1, \tag{5.5b}$$

$$\breve{\boldsymbol{B}} \succeq \mathbf{0}, \tag{5.5c}$$

$$\sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}^c} r_u^c(\mathbf{p}^{c,i}) \xi_i^c \geq R_u + \breve{B}_{uu}(\hat{R} - R_u), \quad \forall u \in \mathcal{U}, \tag{5.5d}$$

$$\text{Constraints in (2.3c) and (2.3d).} \tag{5.5e}$$

A sufficient condition for the Lagrange-dual and the bi-dual of the problem in (5.4) to have equal objectives is the existence of a feasible solution of the bi-dual problem in (5.5) with strictly higher sum-rates than the minimal rates $\mathbf{R}$, i.e., $\breve{B}_{uu} > 0, \forall u \in \mathcal{U}$.

See Appendix A.10 for a proof.

Theorem 11. *We have that $P_{(\mathbf{R},\hat{\mathbf{P}})}^{cov,sdp,ts} \leq P_{(\mathbf{R},\hat{\mathbf{P}})}^{cov,lr,ts}$, i.e., the SDP relaxation is never less tight than the linear relaxation. Furthermore, the conditions $\boldsymbol{b} \succ \mathbf{0}, \sum_{u \in \mathcal{U}} b_u \leq 1$, for the optimal solution in (5.3) are sufficient for $P_{(\mathbf{R},\hat{\mathbf{P}})}^{cov,sdp,ts} = P_{(\mathbf{R},\hat{\mathbf{P}})}^{cov,lr,ts}$ to hold.*

See Appendix A.11 for a proof.

While Theorem 11 indicates that the relaxation in (5.5) can be tighter than the linear relaxation in (5.3), it does not say by how much. To further analyze this question we simulated 1000 randomly generated 4-user DSL scenarios with parameters as specified in Section 5.2.1. We solve the problem in (5.3) to optimality by applying the column generation scheme of Chapter 4, with optimal solutions of the per-subcarrier problems in (4.3b) as described in Chapter 3. After convergence of column generation at iteration $k = \hat{k}$ we reuse the generated power and rate allocations indexed by $i \in \mathcal{I}^{c,(k)}, c \in \mathcal{C}$, for the corresponding SDP relaxation and compare the obtained objectives under both relaxations. We found that the difference between the objectives in relation to that of the continuous relaxation was less than 1 ppm, indicating that the SDP relaxation is not substantially tighter than the continuous relaxation.

In the following section we propose a heuristic for recovering integer solutions $b_u \in \{0, 1\}, u \in \mathcal{U}$, and show further results comparing the two proposed relaxations.

5.1.3 A Heuristic For Coverage Maximization

In the above sections we considered the time-sharing relaxation of the original problem in (5.1) under two further relaxations of the binary service update variables which both yield an estimate $b_u \in [0, 1], u \in \mathcal{U}$, of a binary solution of service

---

Algorithm 6 Coverage Maximization (CM) Heuristic

1: Solve the time-sharing relaxation in (5.3) or (5.5) by column generation, cf. Section 4.1, solving the subproblems in (4.3b) using the heuristics in Section 4.2.
2: Use the obtained variables $b_u \in [0,1], u \in \mathcal{U}$, as priority information in the sequential variable fixing procedure in Section 5.1.3.
3: Resolve another LP as in (5.3) minimizing the sum-power based on the found (fixed) coverage $b_u \in \{0,1\}, u \in \mathcal{U}$, i.e., with target-rates $R_u + b_u(\hat{R} - R_u), u \in \mathcal{U}$.
4: Recover a feasible solution for the problem in (5.1) using the heuristic CHET in Section 4.1.4 with target-rates $R_u + b_u(\hat{R} - R_u), u \in \mathcal{U}$.

---

upgrades for the original problem in (5.1). More precisely, in case of the SDP relaxation we extract this information by taking the eigenvector of the solution $\breve{B}$ according to the largest eigenvalue [157], element-wise divided by the last element. Our next step is a heuristic for sequentially fixing the service variables to either zero or one: We simply try setting the lower-bound of the largest fractional variable $b_u, u = \text{argmax}_{i \in \mathcal{U}}\{b_i\}$, to one and re-solve the master LP in (5.3) from the final column generation iteration $k = \hat{k}$. If the LP is feasible we keep the lower-bound of b_u at one or reset it to zero otherwise, cf. the simple feasibility detection mechanism based on an artificial column in Section 4.1.1. This process is repeated until all variables $b_u, u \in \mathcal{U}$, are fixed at integer values. Next we solve yet another, similar LP with the constraints as above but with lower-bounds of binary variables now set to the found integer values and the objective being the minimization of the total transmit power. This is to avoid unnecessarily wasting power while maintaining the found level of service coverage. Finally we apply the CHET heuristic of Section 4.1.4 to recover from the time-sharing relaxation and to find a feasible solution for our original problem in (5.1), cf. the overview in Algorithm 6. Summarizing, the key ideas of our CM heuristic are a) to exploit the information in the optimized continuous variables $b_u, u \in \mathcal{U}$, in order to prioritize users in the sequential fixing process, and b) to use the memory in column generation based schemes for checking feasibility of specific 0-1 fixings. As a result we obtain a low-complexity method for sequential fixing of fractional variables, where we only re-solve in total $U + 1$ LP's as in (5.3) instead of re-solving $U + 1$ different full DSM problems as in (5.1).

5.2 Performance Evaluation

5.2.1 On the Suboptimality in Small Networks

We will now study the suboptimality of our CM heuristic by comparison to an exhaustive search over all 0-1 combinations of binary variables $b_u, u \in \mathcal{U}$. We randomly generate 2000 downstream, 4-user VDSL [50] scenarios by uniformly sampling the
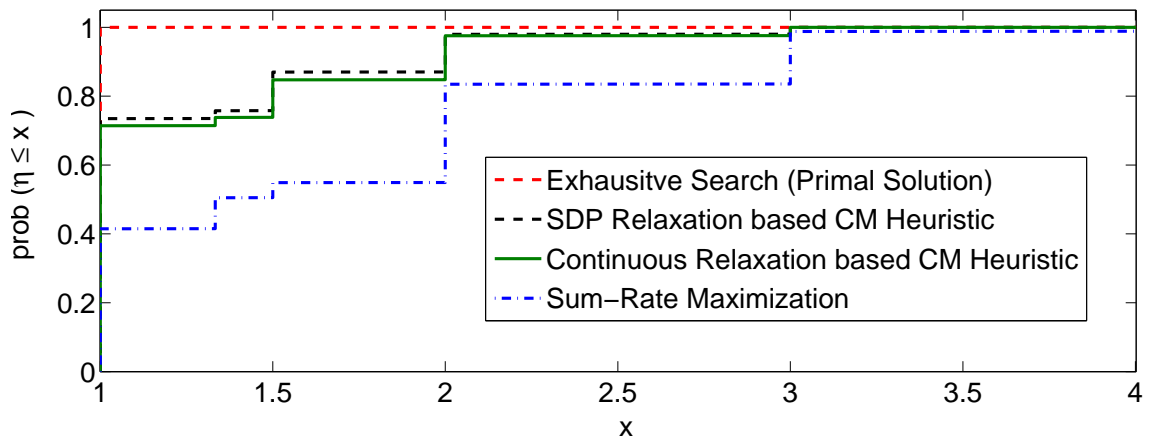
Figure 5.1: Cumulative distribution function of the performance ratio η in (5.6) for various DSM approaches.

lengths $L_{CO-Cab} \in [100, 1400]$ and $L_{Cab-u} \in [50, 500], \forall u \in \mathcal{U}$, cf. Figure 4.5, with 3 and 1 lines deployed from central office and cabinet, respectively. The target-rates are $\hat{R} = 30$ Mbps and $R_u = 6$ Mbps, for all users $u \in \mathcal{U}$, respectively, and the chosen simulation parameters are as detailed in Section 4.2.3. Furthermore, we use the column generation framework of Chapter 4 with optimal solutions of the per-subcarrier problems in (4.3b) as described in Chapter 3 for all of the compared DSM approaches. We define an algorithm's performance ratio η between the optimal objective $P^{cov}_{(\mathbf{R}, \hat{\mathbf{P}})}$ in (5.1) and the algorithm's coverage $P^{\text{cov,Alg}}$ through a feasible solution in (5.1) as

$$\eta = \frac{P^{cov}_{(\mathbf{R}, \hat{\mathbf{P}})}}{P^{\text{cov,Alg}}}. \tag{5.6}$$

Figure 5.1 shows cumulative distributions of the ratio η for our CM heuristic, the primal solution obtained after the exhaustive search through CHET in Section 4.1.4, and the result of a sum-rate maximization. Regarding the upper dashed curve we see that after the exhaustive search a feasible primal solution was obtained in all cases (using CHET), confirming that we truly compare to the *primal* optimum. We find that sum-rate maximization and the CM heuristic (with linear relaxation) are optimal in less than 45 % and more than 68 %, respectively, with 99 % confidence according to a t-test. The performance of the CM heuristic under the SDP relaxation was in various cases seen to be worse than that under the linear relaxation. This is however an artifact of the applied heuristics for recovering binary variables $b_u, u \in \mathcal{U}$, and recovering from time-sharing, cf. Theorem 11. On average the coverage achieved under the SDR relaxation is less than 5 % better than that under the linear relaxation, cf. Figure 5.1. Further results and an intuition behind the good performance of the applied linear relaxation will be developed in in the next section.

5.2.2 Performance Evaluation in Large Networks

We apply our CM algorithm under the linear and SDP relaxations to the problem in (5.1), as well as the sum-rate maximization (RM) scheme in Chapter 4, both using the heuristics rSEGO with a complexity limit of $3 \cdot 10^4$ power evaluations per subcarrier, WS-LS with 40 regular subcarrier intervals shifted in every iteration k by a uniformly drawn random number of subcarriers between zero and one interval length, and other simulation parameters as specified in Section 4.2.3. We randomly generated 25-user DSL networks as described in Section 5.2.1 and assigned 10 users to the cabinet and 15 users to the central office. We compare our CM and RM schemes to optimal single-user bit-loading [27] under worst-case crosstalk noise $\sum_{i \in \mathcal{U} \setminus u} H_{ui}^c \hat{p}_i^c, c \in \mathcal{C}$, ("Mask-Based" SSM) and the sum-rate maximizing algorithms in [103] (MDBL) and a scheme closely related to [31] (ISB).[1] Under RM and CM we assume minimum target-rates $R_u = 6$ Mbps, $\forall u \in \mathcal{U}$. The upgrade target-rate is set to $\hat{R} = 16$ Mbps. Our average simulation results are shown in Figure 5.2 where we observe a large gap in terms of the number of upgraded users between the four tested DSM algorithms and Mask-Based SSM. For example, CM(LR) (that is CM under the linear relaxation) improves upon Mask-Based SSM in this respect by on average more than 82.9 % (8.95 ± 0.41). Comparing CM(LR) to the best-performing rate-maximizing DSM algorithm RM we find that our CM scheme yields a 13.9 % (2.42 ± 0.32 users) increase in service coverage. Intuitively, the improvements by our CM heuristic under the applied continuous relaxation of binary variables in Section 5.1.1 upon the sum-rate maximizing algorithms can be explained by the reduction in rates of the lines which would over-achieve the upgrade target-rate under sum-rate maximization. This results in lower crosstalk into the longer, e.g., central office deployed lines, and consequently in a higher service coverage. Differently to Section 5.2.1 where we applied an optimal algorithm to solve the per-subcarrier subproblems in (4.3b), we see in Figure 5.2 that the SDP relaxation (referred to as "CM (SDP)" in Figure 5.2) performs worse than the linear relaxation under heuristic subproblem solutions.

At the beginning of this chapter we hypothesized that CM is accompanied by a power reduction compared to RM. Figure 5.3 shows the obtained average transmit and line-driver sum-power values in our random scenarios, cf. Section 2.3.1 for the used line-driver model. Comparing RM and CM(LR) we find that CM indeed yields a reduction in transmit and line-driver power of more than 37.5 % (8.88 ± 0.24 mW) and 14.2 % (47.04 ± 0.98 mW), respectively, confirming our intuition.

---

[1] We perform iterative spectrum balancing (ISB) similarly as in [31] but differently using a line-search over bit-rates and a bisection search over Lagrange multipliers. Feasible power-allocations under ISB are obtained by the heuristic in [196]. The used convergence criterion for ISB is a maximum number of 50 user sweeps not improving the Lagrangian or a total of 200 iterations, while we set bounds on the number of iterations for CM and RM as $30 \leq \hat{k} \leq 60$ and stop if the primal improvement over iterations falls below 0.1 ppm. In CM we initially increase $\hat{R}$ and $R_u, u \in \mathcal{U}$, by 3 % but use the original values for the final heuristic CHET.
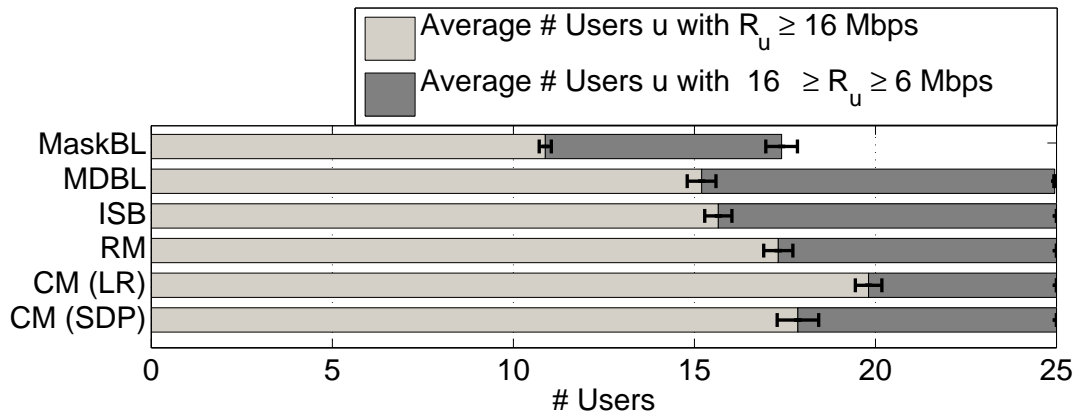
Figure 5.2: Average service coverage in 1000 DSL scenarios.
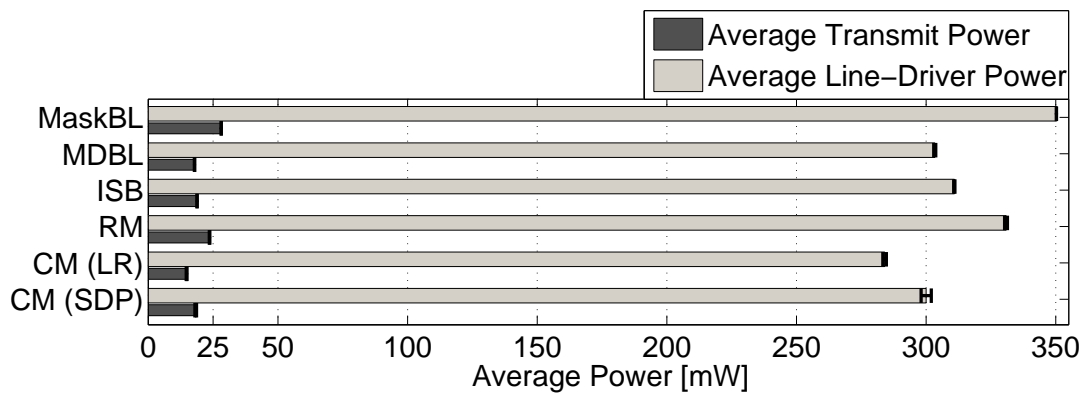
Figure 5.3: Average power consumption in 1000 DSL scenarios.

Part III

DSM Stabilization Techniques

Chapter 6 Robust DSM using Crosstalk Margins

In this chapter we introduce *robust DSM*, an extension to energy-efficient DSM algorithms which is based on margins on the crosstalk couplings and thereby allows to guarantee a minimal rate-performance under a deterministic worst-case crosstalk-channel uncertainty model. While this extension is a contribution in its own right, our main motivation is to conservatively demonstrate the effect of errors in the estimation of the crosstalk gain coefficients, which are a prerequisite for the DSM methods studied in this thesis, on the energy-efficiency achieved by DSM. Minimizing the transmit power based on given service-related target-rates may greatly reduce the power dissipation of a DSL transceiver as shown in Chapters 2 and 5. However, the transmission at minimum power may come at the expense of robustness as the guaranteed service-rates may easily be violated when real and modeled system parameters differ.

We assume the model uncertainties stem from estimation errors. While in current DSL modems the direct channel can be accurately estimated during line initialization, various methods have been proposed in literature to also obtain estimates of the crosstalk channel gains. Non-modem-based NEXT estimation in the frequency domain was proposed in [55] based on correlation with a canonical set of PSD basis-functions that are solely technology dependent. However, this scheme is not applicable to FEXT estimation as FEXT also strongly depends on the loop length. In [201] offline least-squares crosstalk identification is performed based on the recording of transmitted and received signals. In [3] blind crosstalk estimation is performed by expectation maximization. In [128] FEXT estimation is done incrementally, i.e., each active line estimates the coupling from a currently initializing line by estimation using the pseudo-random training sequence of the initializing line. The reverse couplings are obtained similarly by sequentially re-initializing the active lines. The performance is shown to depend on the (unknown) symbol frequency offset, the background noise level, and the number of samples used for least-squares estimation.

Differently to all these approaches, methods which do not necessitate changes in current DSL modems' hardware or software have been proposed in [92, Ch. 3] and [106]. Furthermore, single-ended line testing (SELT) [97] can be used to estimate the loop lengths, which together with the crosstalk model in (1.1) may give a conservative indication of the FEXT gains. Related to the incremental estimation

in [128], the crosstalk gain estimation suggested in [92, Ch. 3] uses the observed variations in SNR feedback when users become inactive / active. This estimate is further improved by combining various estimates and smoothening the estimated frequency response. A practical implementation of DSM based on crosstalk identification using SNR fluctuations is presented in [30]. The least-square estimator in [106] is based on U sequential modem-based single-input, multiple-output PSD measurements. The suggested implementation is by using standardized loop diagnostic features [86]. Estimation errors are reported to be introduced as each PSD measurement is quantized by an integer in dBm/Hz, and due to slight time-variations of the measurements that are for instance attributed to the internal transceiver noise [106]. Laboratory experiments with ADSL2+ modems are carried out and the proposed estimator is found to have a mean deviation of less than $3\,\mathrm{dB}$ from the reference (network-analyzer based) measurements for most frequencies [106]. For our simulations in Section 6.4 we will use this estimation performance as a reference. Various work has studied the impact of crosstalk estimation errors on the rate-performance of DSM [92, Ch. 5], [107], [117]. The specifics of our study compared to these previous ones are that a) we study the impact of crosstalk estimation errors on energy-efficiency, b) we perform DSM already under the assumption of estimation errors, and c) we use a fairly conservative simulation setup. More precisely, we apply the "worst-case" model in (1.1) and additionally assume a deterministic "worst-case" estimation error, cf. Section 6.1 for a precise definition.

A classical approach to describe the uncertainty of optimization parameters is by their probability distributions [94]. Differently, our work is based on the *robust optimization* framework [10, 13], [22, Section 6.4], in which the uncertainty of parameters is modeled by the set of their possible realizations and the optimization is performed with respect to the (a-priori unknown) worst-case element of this set. We motivate the worst-case optimization approach in DSL by arguing that network operators are not interested in a probabilistic long-term outage probability [96] but in a guaranteed minimum rate which is in accordance with the service level promised to the customers. Concluding, for energy-efficient DSM to be accepted by operators it is imperative to explicitly account for model uncertainties during optimization in a worst-case sense.

Examples of how the robust optimization framework has been applied to problems in communications can be found in [1,109,121,122,131,174,193]. Previous approaches to incorporate parameter uncertainty are an adequate noise-margin equally set for all subcarriers as supported by current DSL modems [86], the virtual noise concept [172] which effectively results in arbitrarily set *per-subcarrier* margins, the optimization of outage probabilities [90], or the repeated/iterative optimization triggered by system changes (e.g., online reconfiguration techniques such as "bit-swapping" [86]). While DSM subject to uncertainty in the noise has been studied in [90], it is the uncertainty in crosstalk coefficients we are foremost concerned with. However, an extension of our approach to take uncertainties in the noise-power into account is straightforward.

Uncertainty in channel coefficients may account for channel estimation errors as explained above, as well as implementation errors of the calculated power levels. Implementation errors may occur through the representation of the power-allocation as the interpolation of the power-values at a few frequency sampling-points, or in practical DSM implementations due to the indirect specification of the PSD through the definition of spectral mask constraints [117].

We begin in Section 6.1 by introducing our robust DSM problem formulation. The complexity and conservatism of our worst-case optimization approach depend on the definition of an uncertainty set of the per-subcarrier channel parameters, which can be interpreted as a multi-dimensional margin specifically targeting variations in crosstalk noise. Therefore we motivate in Section 6.2 the use of two tractable convex per-subcarrier uncertainty sets characterized by a few parameters, including the simple idea of scaling crosstalk parameters to their worst-case values (similar to per-subcarrier margins). Furthermore, in Section 6.3 the robust per-subcarrier optimization problems based on these uncertainty sets are shown to be readily integrated into previous DSM schemes. This intuitively holds as the robust per-subcarrier power-allocation problems and the bit-loading problem are in fact separable. One possible application of robust DSM is in "design centering", that is the problem of finding the maximal parameter uncertainty for which a feasible robust solution can be found. Note the similarity of this problem to the common margin-adaptive bit-loading strategy [59]. However, as we target energy-efficiency, in Section 6.4 we will compare the worst-case rate-loss of the optimal *nominal* (i.e., non-robust) DSM solution to the sum-power penalty of robust DSM. Similarly, in [10] the "price of robustness" in linear programming was analyzed in numerous benchmark problems. Our hypothesis is that the power penalty caused by the robust per-subcarrier solutions is controlled by multi-carrier DSM, as the latter allows to assign heavily interfering users to different subcarriers.

6.1 The Robust Spectrum Balancing Concept

We begin our development of energy-efficient robust DSM by introducing uncertainty in the form of a vector $\tilde{\mathbf{H}}_{u:}^c \in \mathcal{R}_+^U$ of *uncertain* crosstalk coefficients to user u on subcarrier c, where $\tilde{H}_{uu}^c = 0$. The rate under these uncertain coefficients is defined similar to that under certain coefficients $\mathbf{H}^c$ in (1.3) as

$$r_u^c(\mathbf{p}^c, \tilde{\mathbf{H}}_{u:}^c) \doteq \log_2\left(1 + \frac{H_{uu}^c p_u^c}{\Gamma(\sum_{i \in \mathcal{U} \setminus u} \tilde{H}_{ui}^c p_i^c + N_u^c)}\right), \tag{6.1}$$

differing from (1.3) only in that the dependency on the crosstalk coefficients $\tilde{\mathbf{H}}_{u:}^c$ is made explicit. Based thereupon we extend the original nominal problem in (1.8) to

its "robust counterpart", referred to as the *primal robust problem*, given as

$$\underset{0 \leq p_u^c \leq \hat{p}_u^c, r_u^c \in \mathcal{B}_u^c, u \in \mathcal{U}, c \in \mathcal{C}}{\text{minimize}} \quad \sum_{c \in \mathcal{C}} f^c(\mathbf{p}_c, \hat{\mathbf{w}}, \check{\mathbf{w}}) \tag{6.2a}$$

$$\text{subject to} \qquad \sum_{c \in \mathcal{C}} r_u^c \geq R_u, \quad \sum_{c \in \mathcal{C}} p_u^c \geq \hat{P}_u, \quad \forall u \in \mathcal{U}, \tag{6.2b}$$

$$r_u^c(\mathbf{p}^c, \tilde{\mathbf{H}}_{u:}^c) \geq r_u^c, \quad \forall u \in \mathcal{U}, c \in \mathcal{C}, \forall \, \tilde{\mathbf{H}}_{u:}^c \in \mathcal{H}_u^c, \tag{6.2c}$$

where $\mathcal{H}_u^c \ni \tilde{\mathbf{H}}_{u:}^c$ denotes the uncertainty set of feasible crosstalk coefficients. Without loss of generality we may restrict ourselves to convex uncertainty sets as stated by the following theorem.

Theorem 12. *The worst-case rate* $\min_{\tilde{\mathbf{H}}_{u:}^c \in \mathcal{H}_u^c} \{r_u^c(\mathbf{p}^c, \tilde{\mathbf{H}}_{u:}^c)\}$ *is invariant when replacing the uncertainty set* $\mathcal{H}_u^c$ *by its convex hull* $\text{conv}(\mathcal{H}_u^c)$.

Proof. First we note the independence among users in terms of uncertainty sets and target-rate constraints, and that $r_u^c(\mathbf{p}^c, \tilde{\mathbf{H}}_{u:}^c)$ in (6.1) is monotonously decreasing in the interference term $\sum_{i \in \mathcal{U} \setminus u} \tilde{H}_{ui}^c p_i^c$. Hence, the worst-case can be constructed by maximizing this interference term. Assuming

$$\tilde{\mathbf{H}}_{u:}^{c,*} = \underset{\tilde{\mathbf{H}}_{u:}^c \in \mathcal{H}_u^c}{\text{argmax}} \left\{ \sum_{i \in \mathcal{U} \setminus u} \tilde{H}_{ui}^c p_i^c \right\}, \tag{6.3}$$

using Carathéodory's theorem [12, Prop. B.6], and remembering that $\tilde{H}_{uu}^c = 0, \forall \tilde{\mathbf{H}}_{u:}^c \in \mathcal{H}_u^c$, we have that

$$\underset{\tilde{\mathbf{H}}_{u:}^c \in \text{conv}(\mathcal{H}_u^c)}{\max} \left\{ \sum_{i \in \mathcal{U} \setminus u} \tilde{H}_{ui}^c p_i^c \right\} = \underset{\substack{\alpha_j \geq 0, \sum_{j \in \mathcal{U}} \alpha_j = 1, \\ \tilde{\mathbf{H}}_{u:}^{c,(j)} \in \mathcal{H}_u^c, \forall j \in \mathcal{U}}}{\max} \left\{ \sum_{i \in \mathcal{U} \setminus u} p_i^c \sum_{j \in \mathcal{U}} \alpha_j \tilde{H}_{ui}^{c,(j)} \right\} \tag{6.4}$$

$$= \underset{\substack{\alpha_j \geq 0, \sum_{j \in \mathcal{U}} \alpha_j = 1, \\ \tilde{\mathbf{H}}_{u:}^{c,(j)} \in \mathcal{H}_u^c, \forall j \in \mathcal{U}}}{\max} \left\{ \sum_{j \in \mathcal{U}} \alpha_j \sum_{i \in \mathcal{U} \setminus u} \tilde{H}_{ui}^{c,(j)} p_i^c \right\} \tag{6.5}$$

$$\leq \sum_{i \in \mathcal{U} \setminus u} \tilde{H}_{ui}^{c,*} p_i^c, \tag{6.6}$$

which concludes the proof. $\square$

As mentioned, our notation implicitly assumes independence of uncertainty among users and subcarriers. The projection of the global uncertainty onto uncertainties per user is natural, since the rate constraints have to be fulfilled for each user separately and an uncertainty correlation among users is irrelevant for worst-case feasibility under parameter uncertainty. The independence assumption among vectors $\tilde{\mathbf{H}}_{u:}^c, \forall c \in \mathcal{C}$, was made to allow an analytic treatment that is independent of the underlying estimation process.
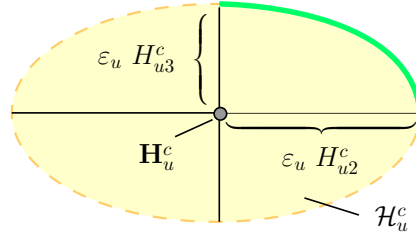
Figure 6.1: Schematic illustration of an ellipsoidal uncertainty region $\mathcal{H}_u^c$ for $u = 1$ and $U = 3$, where the possible interference-maximizing set lying on the relative boundary of $\mathcal{H}_u^c$ is emphasized.

6.2 Uncertainty Regions

In this chapter we have chosen a *multiplicative* model of uncertainty in the crosstalk-coefficients. This choice captures for example arbitrary temporal changes of crosstalk coefficients or an anticipated percental estimation error. A different approach would have been to model uncertainty using an additive error measure. Assuming a norm-constraint on the additive errors one obtains a similar convex subproblem formulation as we will get for ellipsoidal uncertainty, as shown in [193] for single-subcarrier power control. However, to get new insights into the resulting robust problem formulation, and because the measured uncertainty values in [106] which will be of relevance in our later simulations are available in logarithmic scale, we base our derivations on a multiplicative formulation. Furthermore, we may also regard this choice as a way to incorporate worst-case changes in the power-allocation of other users, e.g., due to changes of their direct channel coefficients. Finally we note that qualitatively the specific formulation of uncertainty will not alter the main conclusions we draw from this work.

6.2.1 Ellipsoidal Uncertainty Sets

We will first derive and analyze an uncertainty set in the shape of an ellipsoid. This shape is commonly used due to its relation to Gaussian probability distributions [22]. It was further chosen for its tractability and representativeness for convex uncertainty regions with *interdependence* of worst-case coefficients. Under a multiplicative (percental) uncertainty of the crosstalk coefficients relative to their nominal value, the uncertainty set is given by (cf. Figure 6.1)

$$\mathcal{H}_u^c = \{\tilde{\mathbf{H}}_{u:}^c \, | \tilde{H}_{ui}^c = H_{ui}^c(1 + \Lambda_{ui}^c), \forall i \in \mathcal{U} \setminus \{u\}, \tilde{H}_{uu}^c = 0, \|\mathbf{\Lambda}_{u:}^c\|_2 \leq \varepsilon_u\}, \quad (6.7)$$

where $\mathbf{\Lambda}_{u:}^c \in \mathbb{R}^U$, and $\boldsymbol{\varepsilon} \in \mathbb{R}_+^U$ contains the uncertainty parameters of all users. Note that due to the positivity of the variables p_u^c it is sufficient to restrict our attention to values $\mathbf{\Lambda}_{u:}^c \succeq \mathbf{0}$. A robust per-subcarrier power-allocation subproblem in similar
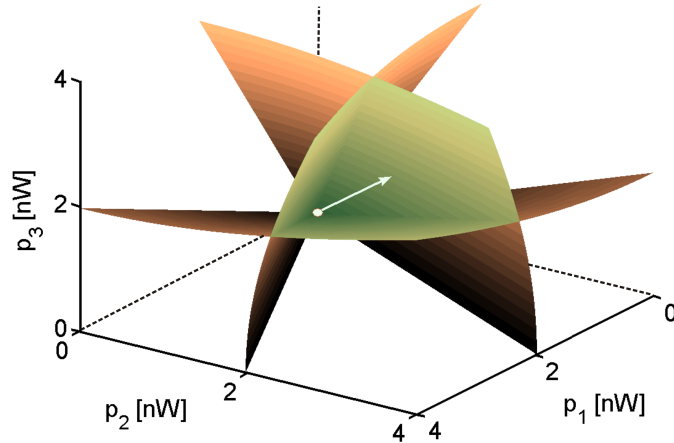
Figure 6.2: Illustration of the boundaries of SNR constraint SOC's and the feasible power region (lightly shaded) for robust single-subcarrier power-allocation with ellipsoid-shaped uncertainty of crosstalk coefficients.

form to the classical power control problem in (1.4) can hence be posed as

$$\underset{0 \leq p_u^c \leq \hat{p}_u^c, \forall u \in \mathcal{U}}{\text{minimize}} \quad \sum_{u \in \mathcal{U}} w_u p_u^c \tag{6.8a}$$

$$\text{subject to} \quad \left(\mathbf{I} - \left(\mathbf{F}^c + \Delta\mathbf{F}^{c,ell}(\mathbf{\Lambda}_{u:}^c)\right)\right)\mathbf{p}^c \succeq \mathbf{n}^c, \tag{6.8b}$$
$$\forall \mathbf{\Lambda}_{u:}^c \in \{\tilde{\mathbf{\Lambda}}_{u:}^c \in \mathbb{R}_+^U \mid \|\tilde{\mathbf{\Lambda}}_{u:}^c\|_2 \leq \varepsilon_u\}, \ \forall u \in \mathcal{U},$$

where $\mathbf{w} \in \mathbb{R}^U$ are constant weights which will be specified in Section 6.3.1, $\Delta\mathbf{F}^{c,ell}(\mathbf{\Lambda}_{u:}^c) \in \mathbb{R}^{U \times U}, \forall c \in \mathcal{C}$, and

$$\Delta F_{uv}^{c,ell}(\mathbf{\Lambda}_{u:}^c) = \begin{cases} 0, & \text{if } v = u, \\ \Gamma\gamma_u^c \Lambda_{uv}^c H_{uv}^c / H_{uu}^c, & \text{otherwise.} \end{cases} \tag{6.9}$$

In this specific case the semi-infinite problem in (6.8) can be cast in a minimax form as shown in [57]. Conferring to the primal robust problem in (6.2), the model-parameters maximizing the interference can be analytically derived using the following relations [22, p. 322]

$$\max_{\tilde{\mathbf{H}}_{u:}^c \in \mathcal{H}_u^c} \left\{ \sum_{i \in \mathcal{U}\setminus u} \tilde{H}_{ui}^c p_i^c \right\} = \sum_{i \in \mathcal{U}\setminus u} H_{ui}^c p_i^c + \max_{\{\mathbf{\Lambda}_{u:}^c; \|\mathbf{\Lambda}_{u:}^c\|_2 \leq \varepsilon_u\}} \left\{ \sum_{i \in \mathcal{U}\setminus u} H_{ui}^c \Lambda_{ui}^c p_i^c \right\} \tag{6.10a}$$

$$= \sum_{i \in \mathcal{U}\setminus u} H_{ui}^c p_i^c + \varepsilon_u \| \text{diag}(\bar{\mathbf{H}}_{u:}^c)\mathbf{p}^c\|_2, \tag{6.10b}$$

where $\text{diag}(\mathbf{x})$ denotes the matrix having the vector $\mathbf{x}$ in the diagonal and zeros elsewhere, and $\bar{\mathbf{H}}_{u:}^c \in \mathbb{R}^U$ is defined as $\bar{H}_{uv}^c = 0$ for $v = u$ and $\bar{H}_{uv}^c = H_{uv}^c$ otherwise.

Hence, the constraint set in (6.8) can be written more compactly as the product-set of second-order cone (SOC) [108] constraints, and the problem in (6.8) is rewritten as

$$\underset{0 \leq p_u^c \leq \hat{p}_u^c, \forall u \in \mathcal{U}}{\text{minimize}} \quad \sum_{u \in \mathcal{U}} w_u p_u^c \tag{6.11a}$$

subject to

$$p_u^c - \Gamma \frac{\gamma_u^c}{H_{uu}^c} \Big(\sum_{i \in \mathcal{U} \backslash u} H_{ui}^c p_i^c + \varepsilon_u \| \operatorname{diag}(\bar{\mathbf{H}}_{u:}^c) \mathbf{p}^c \|_2 \Big) \geq n_u^c, \; \forall u \in \mathcal{U}. \tag{6.11b}$$

Figure 6.2 schematically illustrates a symmetric robust power-allocation problem on subcarrier $c = 1$ for $U = 3$ users each located at $300\,\mathrm{m}$ distance from the deployment point, $\varepsilon_u = 300, r_c^u = 1, \forall u \in \mathcal{U}$, and all other simulation parameters as specified in Section 6.4. Note that the high uncertainty-radius was chosen to emphasize the shape of the constraint set. The lightly shaded area is the intersection of three SOC's as defined by the users' constraints in (6.11b), the arrow represents the constant gradient of the cost to be minimized, and the marked point indicates the optimum.

6.2.2 Box-Shaped Uncertainty Sets

The second type of uncertainty regions we will use are box-shaped ones, given by

$$\mathcal{H}_u^c = \{\tilde{\mathbf{H}}_{u:}^c \,|\, \tilde{H}_{ui}^c = H_{ui}^c(1 + \tilde{\Lambda}_{uv}^c), \forall i \in \mathcal{U} \backslash \{u\}, \tilde{H}_{uu}^c = 0, \tilde{\Lambda}_{uv}^c \leq \varepsilon_u, \forall v \in \mathcal{U}\}, \tag{6.12}$$

where $\tilde{\mathbf{\Lambda}}_{u:}^c \in \mathbb{R}^U$. Note that we use the same uncertainty parameters $\boldsymbol{\varepsilon} \in \mathbb{R}_+^U$ as in the previous section, in order to arrive at a fair comparison to ellipsoidal uncertainty sets and extract the effect of their different shapes. This most pessimistic uncertainty set is equivalent to multiplying the crosstalk coefficients $\mathbf{H}_{u:}^c$ by $(1 + \varepsilon_u), \forall c \in \mathcal{C}, u \in \mathcal{U}$, and can be considered as a margin on the crosstalk noise. The corresponding robust power-allocation problem is posed as

$$\underset{0 \leq p_u^c \leq \hat{p}_u^c, \forall u \in \mathcal{U}}{\text{minimize}} \quad \sum_{u \in \mathcal{U}} w_u p_u^c \tag{6.13a}$$

$$\text{subject to} \quad p_u^c - (1 + \varepsilon_u)\Gamma \frac{\gamma_u^c}{H_{uu}^c} \sum_{i \in \mathcal{U} \backslash u} H_{ui}^c p_i^c \geq n_u^c, \quad \forall u \in \mathcal{U}, \tag{6.13b}$$

which notably has the advantage of remaining a linear program and being solvable by matrix inversion, cf. Section 1.2.3. Comparing the robust subproblems in (6.11) and (6.13), we see that they differ in a single term in the constraint functions. From the triangle inequality, convexity of the square-function, and using $\bar{H}_{uu}^c = 0$, it holds

that

$$\frac{1}{\sqrt{U-1}} \sum_{i \in \mathcal{U} \setminus u} H_{ui}^c p_i^c \leq \| \operatorname{diag}(\bar{\mathbf{H}}_{u:}^c) \mathbf{p}^c \|_2 \leq \sum_{i \in \mathcal{U} \setminus u} H_{ui}^c p_i^c. \tag{6.14}$$

Hence we conclude that the constraint set in (6.11) is never more restrictive than the constraint set in (6.13). Therefore robust power-allocation with box-shaped uncertainty can never yield a lower minimum sum-power than one with ellipsoidal uncertainty region when equal uncertainty parameters are used, which also follows directly by inspection of their respective uncertainty sets in (6.7) and (6.12), respectively.

In [13] a cardinality constrained uncertainty concept was proposed. Therein one assumes box-constraints for each coefficient and restricts the number of coefficients which are allowed to deviate from the nominal value. This approach is usable to model the amount of interference in communication systems where users are entering and leaving the system [193]. Hence, if statistics on the number of users active at any time are available, this approach may be used to ensure robustness in DSL systems when additional lines become active, cf. Chapter 7 where the worst-case crosstalk noise from *all* lines in the system is considered. Note however that the users becoming active then have to adhere to the spectral allocation foreseen/optimized by the DSM algorithm. However, when assuming the maximum of $U-1$ disturbers, one falls back to the box-constrained uncertainty model.

6.2.3 Feasibility Conditions and Outage Probability

Using the effect of uncertainty on the power-allocation constraint in (6.13) together with the uncertainty region in (6.12) and the feasibility criterion in (1.6) for the nominal power-allocation problem in (1.4), it is straightforward to derive the following sufficient feasibility conditions for the robust case with multiplicative uncertainty.

Theorem 13. *A robust, single-subcarrier power-allocation is feasible for the ellipsoidal uncertainty constrained problem in (6.11) and for the box-uncertainty constrained problem in (6.13) with uncertainty parameters $\boldsymbol{\varepsilon} \in \mathcal{R}_+^U$ if*

$$\|\mathbf{F}^c\|_2 < \frac{1}{1 + \sqrt{U}\|\boldsymbol{\varepsilon}\|_2}. \tag{6.15}$$

Furthermore, if $\varepsilon_u = \varepsilon, \forall u \in \mathcal{U}$, then a sufficient condition for ellipsoidal uncertainty is given by

$$\rho\left(\mathbf{F}^c\right) < \frac{1}{1 + \varepsilon}. \tag{6.16}$$

For box-shaped uncertainty sets (6.16) is sufficient and additionally necessary.

See Appendix A.12 for a proof.

We note that the first condition is different to the convergence result in [193] where additive instead of multiplicative uncertainty was assumed, while the latter one turns out equivalent to Lemma 1 in [156] on the problem of power control with SINR margins.

The power-allocation optimization problem in (6.11) for ellipsoidal uncertainty regions can also be cast in a probabilistic framework [22, Section 4.4.2] if we assume independent Gaussian random vectors $\tilde{\mathbf{H}}_u^c, \forall u \in \mathcal{U}, \forall c \in \mathcal{C}$. A probabilistic problem formulation is obtained by replacing ε_u by $\phi^{-1}(1 - \sigma_u)$ and $\mathrm{diag}(\bar{\mathbf{H}}_u^c)$ by $(\boldsymbol{\Sigma}_u^c)^{\frac{1}{2}}$ in (6.11), where σ_u is the desired outage probability, $\boldsymbol{\Sigma}_u^c$ being the covariance matrix for coefficients $\tilde{\mathbf{H}}_u^c$, and $\phi(\cdot)$ denotes the Gaussian cumulative distribution function with zero mean and unit variance.

We are further interested in an outage probability of a robust solution to (6.11) with ellipsoidal uncertainty region when the real uncertainty set is box-shaped (cf. (6.13)), which is *independent* of the exact solutions to (6.11) and (6.13). The following theorem gives loose outage probability bounds under certain assumptions.

Theorem 14. *Assume having coefficients*

$$\tilde{H}_{uv}^c \in \left[H_{uv}^c(1 - \varepsilon_u), H_{uv}^c(1 + \varepsilon_u)\right], \quad u \in \mathcal{U}, v \in \mathcal{U} \setminus \{u\}, \tag{6.17}$$

all being independent and symmetrically distributed in the interval. The outage probabilities $\sigma_u = \sigma, \forall u \in \mathcal{U}$, according to a robust solution of Problem (6.11) are bounded by

$$\sigma \leq \frac{1}{\sqrt{e}}, \tag{6.18}$$

independently of the scenario, uncertainty parameters ε_u, and the solution $\mathbf{p}^{c,*}$, where e is Euler's number. Another bound of the outage probability is obtained by restriction to the case of independent and uniformly distributed coefficients $\tilde{H}_{uv}^c$ as

$$\sigma = 1 - \frac{V^{ball}(U - 1)}{V^{box}(U - 1)}, \tag{6.19}$$

where $V^{ball}(d)$ and $V^{box}(d)$ are the volumes of the unit ball and the unit box in d dimensions.

See Appendix A.13 for a proof.

We note that the bound in (6.19) for uniform distributions is monotonously increasing in U. Related to this, in the results of Section 6.4.2 we will observe a larger performance gap between the two uncertainty sets for networks with a higher number of lines. Eventually the bound in (6.19) becomes looser than the more general bound in (6.18) when $U \geq 4$.

Algorithm 7 Lagrange Relaxation based Robust DSM

1: **while** Master Problem (Maximization in (1.10)) not Solved **do**
2:   Generate a New Set of Dual Variables $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ (e.g.,
     By the LP based Column-Generation Scheme in Section 4.1.1)
3:   **for** All Subcarriers $c \in \mathcal{C}$ **do**
4:     **while** Optimal Bit- and Power-Allocation Not Found **do**
5:       Follow an optimal, discrete search method (e.g., the BnB Search in Section 3.1.2.1) to Obtain Another $\mathbf{r}^c$
6:       Evaluate $\mathbf{p}^c(\mathbf{r}^c)$ and the Objective $f^c(\mathbf{p}^c, \hat{\mathbf{w}}+\boldsymbol{\nu}, \breve{\mathbf{w}}+\boldsymbol{\lambda})$ in (1.12) by Solving the LP in (6.13) or the SOC problem in (6.11) with weights $\mathbf{w} = (\hat{\mathbf{w}}+\boldsymbol{\nu})$
7:     **end while**
8:     *Optional* (for the specific multiplier search scheme in Section 4.1.1, cf. Line 2): Recover a solution to the problem in (1.8) by the Heuristic in Algorithm 2
9:   **end for**
10: **end while**

6.3 Integration of Robustness in DSM

6.3.1 A Robust DSM Algorithm

Our robust DSM approach can be readily integrated into Lagrange relaxation based DSM algorithms, as exemplified in Algorithm 7. More precisely, we use the column generation scheme of Section 4.1.1 to optimize the dual variables, and the depth-first BnB search of Section 3.1.2.1, notably without the search-space reduction scheme of Section 3.2 which was derived for the non-robust power control problem in (1.4). Line 6 makes the DSM algorithm worst-case robust against uncertainties as specified in Section 6.2. We use once more the heuristic of Section 4.1.4 in Line 10 of Algorithm 1 to return to a feasible solution to our original robust DSM problem in (6.2).

6.3.2 Complexity and Distributiveness

For each evaluated bit-allocation we face the subproblem of computing a robust power-allocation in Line 6 of Algorithm 7. The latter has in the case of a SOC-problem [108] as well as in case of a matrix inversion an asymptotic complexity of $O(U^3)$, assuming the matrix inversion for the case of box-uncertainty is performed by Gaussian elimination. In practice the CPU time to solve the respective subproblems however is to our experience several magnitudes higher in case of ellipsoidal uncertainty.

As the dual optimal scheme in Algorithm 7 is only practical for a very moderate number of users, it is also interesting to see how the two approaches are imple-

mentable in schemes with lower complexity. For example, the heuristics of Section 4.2 are extendable in a straightforward fashion by evaluating the objective as done in Line 6 of Algorithm 7. Various other DSM schemes such as ISB [31] apply sequential iterations over users to optimize the power-allocation, cf. [184] for an application to a sum-power minimization problem. Assuming a fixed power-allocation of all users $i \in \mathcal{U} \setminus \{u\}$, the power p_u^c necessary for given SINR γ_u^c of user u is given as

$$p_u^c = \Gamma \frac{\gamma_u^c}{H_{uu}^c} \Big(\sum_{i \in \mathcal{U} \setminus u} H_{ui}^c p_i^c + \varepsilon_u \| \operatorname{diag}(\bar{\mathbf{H}}_{u:}^c) \mathbf{p}^c \|_2 \Big) + n_u^c, \qquad (6.20)$$

where notably the right hand side does not depend on p_u^c, cf. the problem in (6.11). A similar expression of the minimal power-allocation for a given bit-allocation can be derived from the robust subproblem in (6.13). However, in both cases the resulting power-allocation needs to be compared to the spectral mask in order to judge feasibility of the rate allocation. Hence, as the scheme in [31] foresees an exhaustive search of all bit-allocations $r_u^c \in \mathcal{B}$, we recognize that both approaches suddenly have a similar complexity for evaluating the robust power-allocation.

Besides computational complexity, another practical point is the possibility of distributed implementations. The continuous and heuristic energy-efficient spectrum balancing algorithm in [186], a modification of the original rate-maximizing scheme in [129], theoretically allows for a semi-distributed implementation based on modems measuring their total noise and a spectrum management center (SMC) having information about cross-channel couplings . This message-exchange scheme naturally carries over to box-constrained uncertainty, where each modem u additionally has to measure the background noise separately to derive the pure crosstalk noise and the extra virtual noise term $(1+\varepsilon_u)\frac{\Gamma}{H_{uu}^c}(\sum_{i \in \mathcal{U} \setminus u} H_{ui}^c p_i^c + N_u^c)$, cf. (6.13), and appropriately scales the feedback messages to the SMC. Similarly, an iterative water-filling type of algorithm [184] can be deployed under box-shaped uncertainty. Unfortunately there appears to be no way to efficiently derive or measure the term $\frac{1}{H_{uu}^c}\| \operatorname{diag}(\bar{\mathbf{H}}_{u:}^c) \mathbf{p}^c \|_2$ distributively, which is why both distributed schemes, the heuristic in [186] as well as iterative water-filling, lose their attractiveness under the ellipsoidal uncertainty approach (or in fact *any* other non-box-shaped uncertainty region).

Hence we see that the SOC-constraints in (6.11) couple the power-allocations of all users on each subcarrier. The distributed robust power-control algorithm for additive, ellipsoid-shaped uncertainty in [193] allows for delayed messaging of the coupling terms. Still, in a multi-carrier system this means that a user has to transmit the whole optimized power-allocation $p_u^c, \forall c \in \mathcal{C}$, after his iteration in order for other users to be able to make their decisions distributively, which decreases the attractiveness of distributed DSM schemes.

6.4 Simulation Results

The aim of the following simulations[1] is to evaluate and compare the two proposed types of crosstalk margin in terms of the extra energy consumed and the loss in rate. We consider VDSL downstream scenarios, where 3 or 33 users are equally distributed at $200\,\mathrm{m}$, $400\,\mathrm{m}$, and $600\,\mathrm{m}$ distance from the the deployment point, respectively. We note that there is no difference between ellipsoidal and box-shaped uncertainty regions when only two users are considered. However, the consideration of a 3 user scenario already allows to study the differences between the considered uncertainty regions (e.g., using Algorithm 7), while keeping the combinatorial complexity at a manageable level. In the case of 33 users we assume an equal bit- and power-allocation of all collocated lines, which effectively results in the optimization of 3 user groups and which can be approached by a dual optimal DSM algorithm as in Algorithm 7, cf. Section 7.2.4 for further details. For ease of exposition, we assume an identical uncertainty parameter for all users, i.e., $\varepsilon_u = \varepsilon, \forall u \in \mathcal{U}$. The values for the uncertainty parameters ε will be varied around the value one, based on real-world crosstalk parameter uncertainties indicated in [106].

6.4.1 Robust Sum-Rate Maximization

Figure 6.3 illustrates the optimal power density spectrum of the longest ($600\,\mathrm{m}$ long) lines for an uncertainty parameter $\varepsilon = 5$ (which under box-shaped uncertainty sets is equivalent to a crosstalk noise level virtually increased by approximately $7.8\,\mathrm{dB}$) in a sum-rate maximization problem with 33 lines, using an equal power-allocation for collocated lines. As intuitively expected in this downstream scenario, the bit-allocation of the longest lines tends to avoid crosstalk when a robust power-allocation is used. Hence, especially the spectrum at higher frequencies where the crosstalk into the shorter lines is stronger is then partially abandoned by the longest lines.

In Figure 6.4 we compare the mean rate obtained by robust sum-rate maximization to the mean rate achieved by the nominal sum-rate maximizing solution (power-allocation) under worst-case crosstalk coefficients. More precisely, the (continuous) bit-allocation for the nominal solution is computed from the nominal power-allocation and assuming the worst-case channel coefficients on each subcarrier taken from the ellipsoidal uncertainty set. The worst-case rates for the nominal solution are either computed as a) the rates based on the resulting worst-case continuous bit-allocation, or b) the rates obtained after flooring the worst-case bit-allocation. The intuition is that these two cases indicate the range of rates in which the nominal power-allocation would result for the worst-case channel, having in mind that in

---

[1] The simulation parameters were chosen according to the VDSL standard in [50], with an SNR gap $\Gamma = 12.8\,\mathrm{dB}$ and two transmission bands as defined in band plan 997. The background noise comprised VDSL noise A [50] added to a constant noise floor at $-140\,\mathrm{dBm/Hz}$.
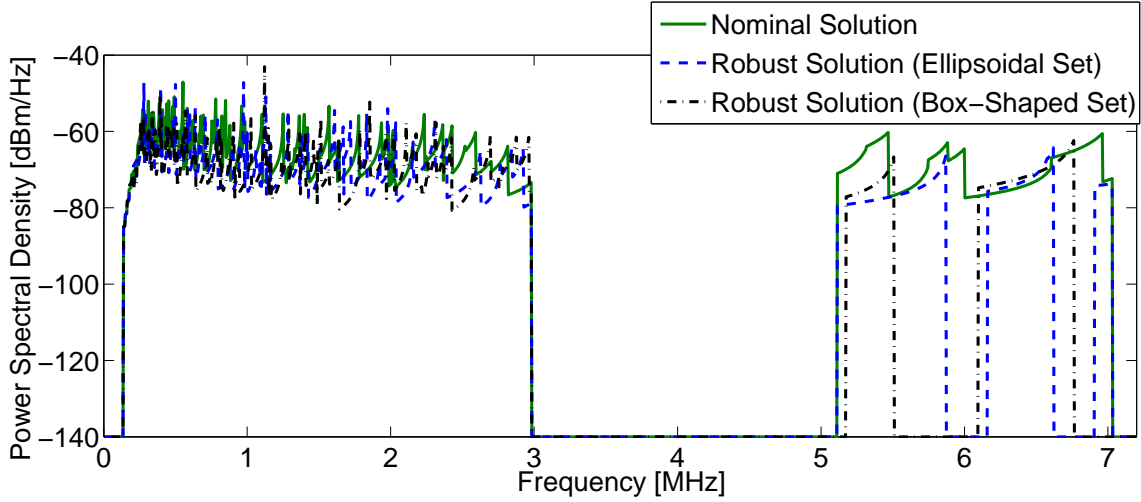
Figure 6.3: Nominal and robust power spectra of the longest lines at the sum-rate maximizing solutions for $\varepsilon = 5$ and 33 lines.
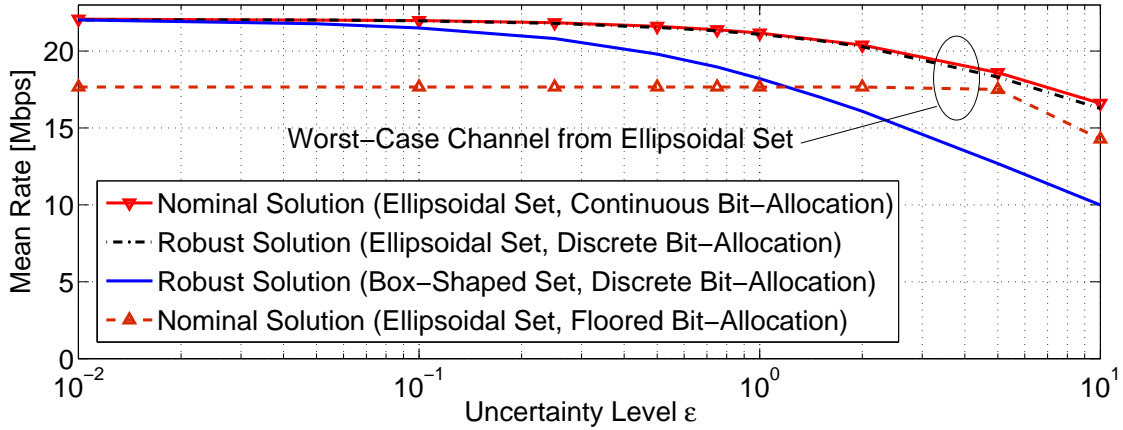
Figure 6.4: Comparison of the robust and nominal solutions in terms of mean rate per user under the worst-case channel coefficients.

practice there are various mechanisms available which let the modems autonomously adapt the bit-allocation when too-high bit-error rates are perceived [59, Sec. 7.5]. While we find that the mean rate under ellipsoidal uncertainty is fairly close to the nominal one, there is a large gap under the box-shaped uncertainty. This can be explained by the fairly large number of lines and the increasing ratio between the volumes of a box and the enclosed ellipsoid with increasing dimension, as already noted in Section 6.2.3. Intuitively, the ellipsoidal uncertainty set becomes less and less conservative when the number of disturbers increases.
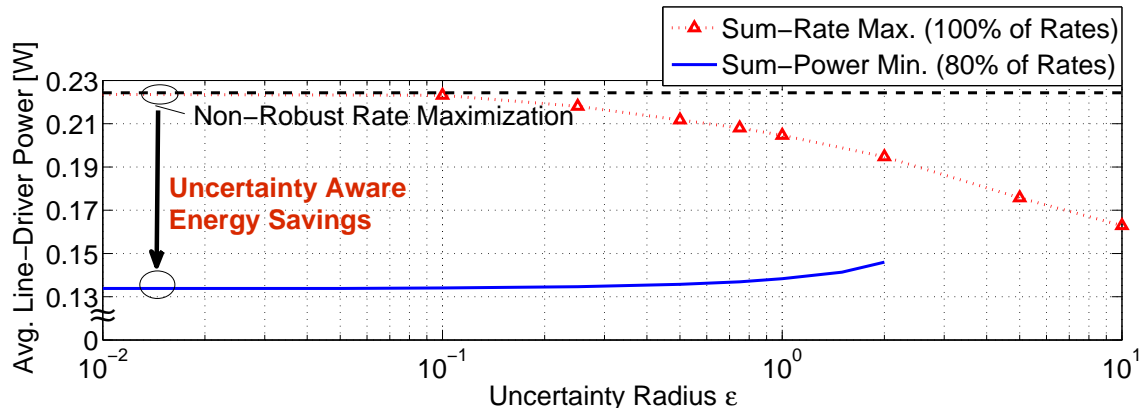
Figure 6.5: Line-driver power savings by robust power minimization compared to a non-robust sum-rate maximization in a 3-user scenario.

6.4.2 Energy Cost for Robustness

One motivation for robust optimization is to guarantee a stable target-rate when these are intentionally reduced to save energy. However, the extra robustness we gain by robust DSM will certainly have a price in terms of a higher power consumption (the "price of robustness" [13]). Hence, we investigate the relation between saved power by rate reduction and extra spent power by robust DSM.

In Figure 6.5 we show the average line-driver power[2] consumption for 3 lines obtained at the non-robust maximum sum-rate solution and a robust power minimization solution with box-shaped uncertainty set and targeting 80 % of the maximum non-robust rates per line. While the power gains by energy minimization compared to the robust sum-rate maximization decrease with an increasing uncertainty, compared to the non-robust solution we find a substantial power saving potential by robust energy minimization.

In Figure 6.6 we show simulation results for a more realistic scenario with three distributed groups of lines and 11 lines each. Due to the larger number of users compared to the 3-user case we find that the (e.g., non-robust) sum-rate maximizing DSM solution already leads to a reduced power consumption. Furthermore, we have already seen in Section 2.3.4 that VDSL lines have a lower potential for power saving compared to ADSL lines. However, Figure 6.6 does show once more that robust solutions still allow to harvest the majority of the possible power savings that would be achievable by transmitting at zero power.

In [188] we demonstrate the minimal transmit sum-power as a function of the uncertainty parameter ε in a distributed 3-user upstream scenario. These results show that at high uncertainty values ε the solutions under both uncertainty sets reduce

---

[2]Parameters of the class-AB line-driver model for VDSL are, similarly as for Section 2.3, given in the simulation section in [185].
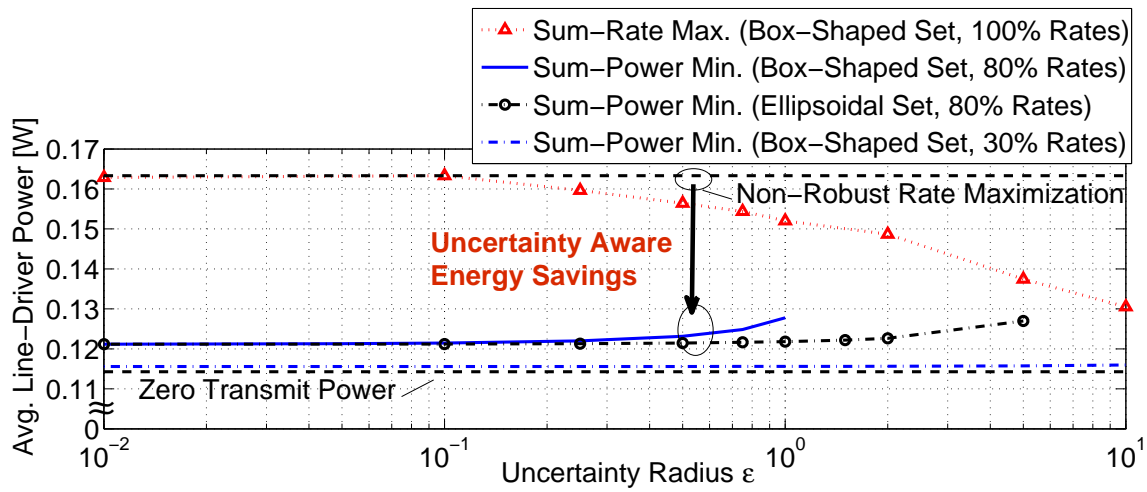
Figure 6.6: Line-driver power savings by robust power minimization compared to a non-robust sum-rate maximization in a 33-user scenario.

the frequency-overlap between the users and therefore give similar transmit powers at high uncertainties. Under the setup of this section the solutions for collocated lines are identical and the crosstalk scenario more severe due to the larger number of DSL lines. Consequently we find a notable difference between the uncertainty sets in terms of the range of uncertainties ε for which feasible solutions were found, cf. Figure 6.6.

Chapter 7 DSM for the Stabilization of DSL under Varying Crosstalk Noise

The activation of low-power modes (LPM) in the DSL access network implies frequent transmit-power changes, resulting in SNR variations with which current DSL systems can hardly cope, cf. Section 2.4 for an introduction to LPMs in DSL. We study the problem of stabilizing DSL systems when using LPMs. A solution proposed in [69] is to inject "artificial noise" (AN) at the transmitter in order to shift eventual variations of the crosstalk noise inside the specified SNR margin which is considered during the initialization of the modems. Besides the empirical description in [69], to the best of our knowledge the problem of setting the AN power spectrum has not been studied so far. Our contribution is the solution of various optimization problems related to the performance evaluation of AN-enabled DSL networks, considering also regulatory and system constraints. Under a worst-case crosstalk noise assumption similar to that made in practice [167] we demonstrate an analytical AN power solution and show the gain by jointly optimizing the AN power spectrum and the SNR margin. However, we embed the problem in a more general mathematical programming framework, which in addition to this worst-case optimization allows us to derive a tight performance upper-bound for AN-enabled DSL systems under specific assumptions. Related to the computation of this bound we find that the optimization of AN seamlessly integrates into previous DSM algorithms and show *provably* near-optimal results for this bound. Our key messages are that a) the performance under AN heavily depends on the set SNR margin and hence AN needs to be optimized jointly with the margin; and b) the injected AN improves the stable rate and power levels compared to systems which are stabilized by SNR margins only, but shows a performance loss for longer loops (e.g., above 1.5 km) compared to an "ideal", frequency selective margin setting (commonly referred to as "virtual noise" [172]).

Background Information and Outline

The activation of LPMs in DSL has already been motivated in Section 2.4 by the achievable energy-savings at the central office (CO). However, the main concern

brought up against LPMs is the fluctuation in the crosstalk noise received on a line when other users enter or leave LPMs. This may lead to an increase in bit-error rate which forces the modems to reinitialize. The two proposals in [47] to resolve this issue partly require changes in standards or do not allow for immediate transitions back into the full-power mode (e.g., in less than 3 s as foreseen in [86]). The most promising proposal made in [69] is to physically inject additional "artificial noise" (AN) at the transmitter (i.e., at the central office) and thereby to reduce the relative impact of future crosstalk variations on the perceived SNR. A related technique is "virtual noise" (VN) [172] which is a frequency selective, tunable receiver-noise parameter masking the crosstalk noise. However, VN is currently not standardized for ADSL systems. Differently to VN, the AN is not masking the changing noise scenario but leads to additional received noise which makes these disturbances fit inside the used SNR margin. In other words, increasing the background noise decreases the SNR reduction when further crosstalk noise is added on top. This approach does not need any standardization effort and supports fast power state transitions. The only theoretical disadvantage is that the AN leads to a higher transmit power, background noise, and crosstalk noise levels. This may reduce the achievable bit-rate and conflict with our initial intention behind the usage of AN: namely to enable LPMs and thereby to reduce the energy consumption in DSL. In [41] it is argued that spectrum balancing in combination with bit-swapping [173] is the more energy-efficient solution for current networks compared to VN, and hence also compared to AN. However, this assumes the availability of the corresponding features in the modems. Furthermore, we deem the majority of possible power savings in DSL coming from the intensive usage of LPMs, which incurs much stronger crosstalk fluctuations than occurring in current networks.

We begin in Section 7.1 by defining the formal system constraints and theoretical multi-user optimization problem for jointly setting the AN power spectra, the transmit-power spectra, and the bit-allocation. This multi-user formulation is notably more general than actually needed for the single-user, worst-case crosstalk noise problem faced in practice [167]: that is the stabilization of a line for the worst-case crosstalk noise. However, the observed "worst-case" noise during the initialization of a DSL connection depends not only on the channel but, for example, also on the lines' target-rates, SNR margins, the user behavior (line usage), and the sequence in which the modems are activated. In order to facilitate a deterministic performance evaluation of AN-enabled networks we study three approaches: a) the joint optimization of AN with the transmit-power spectra, bit-allocations, and SNR margins, b) the single-user bit-loading problem stabilizing the line for the worst-case crosstalk noise, and c) the multi-user sequential initialization of the lines assuming the same worst-case noise as in b) but considering the actual crosstalk noise levels at initialization. Approach a) effectively allows to compute an optimization based bound on the performance under approach b), and is studied in Section 7.2. In this theoretical setting we further investigate in Section 7.2.3 the impact of setting dif-
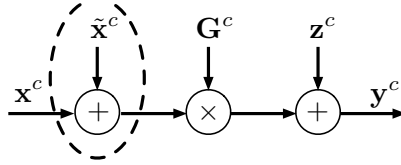
Figure 7.1: Downstream signal model for the U-user interference channel on subcarrier c with all users' received signals $\mathbf{y}^c \in \mathbb{C}^U$, channel matrix $\mathbf{G}^c \in \mathbb{C}^{U \times U}$, and additive noise $\mathbf{z}^c \in \mathbb{C}^U$, showing the addition of the artificial noise $\tilde{\mathbf{x}}^c \in \mathbb{C}^U$ to the transmitted symbols $\mathbf{x}^c \in \mathbb{C}^U$.

ferent SNR margins for different lines by means of a novel margin-search heuristic. Note that this heuristic has also applicability in networks which are not AN-enabled. Similarly, the simplified DSM approach for large networks in Section 7.2.5, based on the assumption of an identical spectral power-allocation for collocated lines, is applicable to the performance evaluation in networks with and without AN capabilities. In Section 7.3 we study approaches b) and c), and derive an analytical solution for the AN power spectrum and an optimal bit-loading algorithm for AN-enabled networks. The performance under the proposed approaches is compared and the value of explicit SNR margin optimization demonstrated by simulations.

7.1 Optimization Problem with Artificial Noise (AN)

7.1.1 Modified System Model

We extend the interference channel model on each subcarrier $c \in \mathcal{C}$ and assumptions of Section 1.2.3 by an independent complex random signal (the "artificial noise") $\tilde{\mathbf{x}}^c \in \mathbb{C}^U$, $\tilde{x}_u^c \sim \mathcal{N}(0, a_u^c)$, being added to the transmitted symbols $\mathbf{x}^c \in \mathbb{C}^U$, $x_u^c \sim \mathcal{N}(0, p_u^c)$ of all users $u \in \mathcal{U}$, cf. Figure 7.1, where $\mathbb{C}^U$ means the set of complex U-dimensional vectors. Similarly as in (1.3) we obtain the achievable rate under a given bit-error probability as [59]

$$r_u^c (\mathbf{p}^c, \mathbf{a}^c) = \log_2 \left(1 + \frac{H_{uu}^c p_u^c}{\Gamma \left(\sum_{i \in \mathcal{U} \setminus u} H_{ui}^c (p_i^c + a_i^c) + H_{uu}^c a_u^c + N_u^c \right)} \right), \qquad (7.1)$$

where $\mathbf{a}^c \in \mathbb{R}^U$ are the AN power levels on all lines. As can be seen in Figure 7.1 and in (7.1) the added AN also generates additional crosstalk noise $\sum_{i \in \mathcal{U} \setminus u} H_{ui}^c a_i^c$ among the users. The reverse, i.e., the powers $\mathbf{p}^c(\mathbf{r}^c)$ for given rates $\mathbf{r}^c$ and constant AN $\mathbf{a}^c$, $c \in \mathcal{C}$, can be computed by the solution of the linear matrix equation in (1.5). However, AN is variable and should be chosen such that the worst SNR a user may experience with respect to crosstalk noise (i.e., when all other users are

transmitting) does not exceed the SNR which was targeted at the initialization of the line. This constraint can be precisely written as

$$\tilde{r}_u^c \left(p_u^c, a_u^c \right) \leq r_u^c \left(\mathbf{p}^c, \mathbf{a}^c \right), \tag{7.2}$$

where $\tilde{r}_u^c \left(p_u^c, a_u^c \right)$ denotes the rate at initialization. The rate $\tilde{r}_u^c \left(p_u^c, a_u^c \right)$ will be the largest, and therefore the constraint in (7.2) the most restrictive, if no crosstalk noise is present at the initialization phase. Hence we define

$$\tilde{r}_u^c \left(p_u^c, a_u^c \right) = \log_2 \left(1 + \frac{H_{uu}^c p_u^c}{\gamma_u \Gamma \left(N_u^c + H_{uu}^c a_u^c \right)} \right) \tag{7.3}$$

where γ_u is the extra SNR margin which is set at the line initialization phase and used for protection against fluctuations in crosstalk noise.[1] Recommended margin values used in practice are in the range between $6 \, \text{dB}$ and $10 \, \text{dB}$ [195]. However, in later sections we will optimize the margin value specifically for the considered networks.

7.1.2 Formulation for the Optimization of Artificial Noise

By extending the original problem in (1.2.4) and considering the constraints of Sections 1.2.3 and 7.1.1, the problem of jointly optimizing transmit power and artificial noise can be cast as

$$\underset{\mathbf{p}^c \succeq \mathbf{0}, \mathbf{a}^c \succeq \mathbf{0}, c \in \mathcal{C}}{\text{minimize}} \quad \sum_{c \in \mathcal{C}} f^c(\mathbf{p}^c, \mathbf{a}^c, \hat{\mathbf{w}}, \check{\mathbf{w}}) \tag{7.4a}$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} \tilde{r}_u^c \left(\mathbf{p}^c \right) \geq R_u, \quad \sum_{c \in \mathcal{C}} (p_u^c + a_u^c) \leq \hat{P}_u, \quad \forall u \in \mathcal{U}, \tag{7.4b}$$

$$\tilde{r}_u^c \left(\mathbf{p}^c \right) \in \mathcal{B}, \quad (p_u^c + a_u^c) \leq \hat{p}_u^c, \quad \forall c \in \mathcal{C}, u \in \mathcal{U}, \tag{7.4c}$$

$$\tilde{r}_u^c \left(p_u^c, a_u^c \right) \leq r_u^c \left(\mathbf{p}^c, \mathbf{a}^c \right), \quad \forall c \in \mathcal{C}, u \in \mathcal{U}. \tag{7.4d}$$

where the objective is similarly as in (1.9) defined as

$$f^c(\mathbf{p}^c, \mathbf{a}^c, \hat{\mathbf{w}}, \check{\mathbf{w}}) = \sum_{u \in \mathcal{U}} \left(\hat{w}_u \left(p_u^c + a_u^c \right) - \check{w}_u \tilde{r}_u^c(\mathbf{p}^c, \mathbf{a}^c) \right). \tag{7.5}$$

We emphasize once more that the purpose of formulating this problem is to derive a performance bound for AN-enabled DSL networks which can be compared to other stabilization techniques. Before presenting an algorithm for the solution of this problem in Section 7.2 we proceed by studying one of its essential building

---

[1] While in all other chapters we assume that the SNR gap Γ already includes the margin γ_u, in this chapter we explicitly state the margin value as we study the performance under varying crosstalk noise, which is part of the disturbance the margin is supposed to protect against.

blocks, that is the per-subcarrier bit and power allocation-problem.

7.1.3 Stabilized Power Control with Artificial Noise

The classical single-subcarrier power control problem [54, 194] for fixed, minimum bit-allocation $\mathbf{r}^c$ can be cast as the linear program [14] (LP) in (1.4) and its solution $\mathbf{p}^c(\mathbf{r}^c)$ obtained by solving the linear system in (1.5). Similarly, we will show that the single-subcarrier power control problem including AN remains an LP. The value of this observation is that discrete-rate DSM algorithms as in [163] or Part II, which rely on the solution of a series of such power control problems, can be extended in a straightforward way to cover also the optimization of the AN, cf. Section 7.2 for details. Regarding only the per-subcarrier constraints in (7.4c) – (7.4d) we can write the power control problem on subcarrier $c \in \mathcal{C}$ for the joint optimization of transmit power and AN power under a fixed bit-load $\mathbf{r}^c \in \mathbb{R}^U$ as

$$\underset{p_u^c \geq 0, a_u^c, \forall u \in \mathcal{U}}{\text{minimize}} \ \sum_{u \in \mathcal{U}} w_u(p_u^c + a_u^c) \tag{7.6a}$$

$$\text{subject to} \ \ H_{uu}^c p_u^c \geq (2^{r_u^c} - 1)\gamma_u \Gamma \big(N_u^c + H_{uu}^c a_u^c\big), \ \ \forall u \in \mathcal{U}, \tag{7.6b}$$

$$\sum_{i \in \mathcal{U} \backslash u} H_{ui}^c \, (p_i^c + a_i^c) + H_{uu}^c a_u^c + N_u^c \leq \gamma_u \big(N_u^c + H_{uu}^c a_u^c\big),$$

$$\forall (c, u) \in \{(c \in \mathcal{C}, u \in \mathcal{U}) | r_u^c > 0\}, \tag{7.6c}$$

$$(p_u^c + a_u^c) \leq \hat{p}_u^c, \quad \forall u \in \mathcal{U}, \tag{7.6d}$$

where $\mathbf{w} = (\hat{\mathbf{w}} + \boldsymbol{\nu})$, $\boldsymbol{\nu} \in \mathbb{R}^U$ are additional weights which will be specified in Section 7.2, and the initial bit-loading constraint in (7.6b) is simply a reformulation of the constraint $\tilde{r}_u^c \, (p_u^c, a_u^c) \geq r_u^c$ using (7.3). Moreover, the SNR variation constraint in (7.6c) is a reformulation of (7.4d) using the rate-definitions in (7.1) and (7.3). The fact that in (7.6c) we possibly only constrain a subset of the users comes from the observation that the constraint in (7.4d) is only active when a user u transmits on subcarrier c. Furthermore, in this case the constraint in (7.4d) only restricts the denominators of the SNR terms in (7.1) and (7.3), cf. (7.6c). Note that $\gamma_u > 1$ has to hold strictly in order for (7.6c) to be feasible under non-zero crosstalk noise and bit-load, cf. (7.6c). The problem in (7.6) is an efficiently solvable LP [14].

7.1.4 Stabilized Power Control without Extra Noise

Another "stabilized" power control problem, which is below referred to as "Margin Only", is readily defined by dropping the AN terms in (7.6) as

$$\underset{0 \leq p_u^c \leq \hat{p}_u^c, \forall u \in \mathcal{U}}{\text{minimize}} \quad \sum_{u \in \mathcal{U}} w_u p_u^c \tag{7.7a}$$

$$\text{subject to } H_{uu}^c p_u^c \geq (2^{r_u^c} - 1)\gamma_u \Gamma N_u^c, \ \forall u \in \mathcal{U}, \tag{7.7b}$$

$$\sum_{i \in \mathcal{U} \setminus u} H_{ui}^c p_i^c + N_u^c \leq \gamma_u N_u^c, \quad \forall (c, u) \in \{(c \in \mathcal{C}, u \in \mathcal{U}) | r_u^c > 0\}, \tag{7.7c}$$

In this formulation the stabilization comes solely from the selective bit-allocation, as can be seen regarding the constraint in (7.7c) which limits the total crosstalk noise received on each line. The solution of the problem in (7.7) can be given analytically, as the constraint in (7.7b), when changed to an equality, provides us with the lowest power values p_u^c, independently for each user $u \in \mathcal{U}$. All that remains to be done is to evaluate feasibility of these values for the constraints in (7.7c) and the power mask constraints in (7.7a) which are the loosest for the found smallest values of p_u^c. The optimal objective can then be evaluated in (7.7a), or assigned infinity in case of infeasibility.

7.1.5 Stabilized Power Control with Virtual Noise

As mentioned in the introduction, the "ideal" alternative to AN, standardized for VDSL2 systems [88], is "virtual noise" (VN) [172]. VN has a similar effect as AN but is only a transmission parameter, i.e., not physically present on the line. The optimal "receiver-referred" VN equals the crosstalk noise spectrum a line experiences when all lines are active, and is hence computable by current DSM schemes [33, 163, 178]. More precisely, one first calculates the optimized received crosstalk noise levels by multi-user DSM (setting the redundant margin to $\gamma = 1$) and then sets $v_u^c = \sum_{i \in \mathcal{U} \setminus u} H_{ui}^c p_i^c$, where $v_u^c, u \in \mathcal{U}, c \in \mathcal{C}$, is the received VN power level considered at the initialization and added to the background noise N_u^c. For comparison to AN we will however also study the (only theoretically relevant) case where $\gamma_u > 1$ and we enforce $v_u^c \geq 0$. In this case we need to solve linear subproblems similarly as in (7.7), only differing in the variable VN terms v_u^c which are added to the background noise in (7.7b) and to the background noise on the right-hand side of the inequality in (7.7c). Another definition of VN which is used in standardization [88] and which will be used in later sections comes from referring to it as a transmit-power level, giving the "transmitter-referred" VN v_u^c/H_{uu}^c. However, this VN parameter is simply a scaled variant of v_u^c, and hence the way to compute the VN mentioned before remains valid.

After having studied the per-subcarrier power control problem for the cases with AN, with VN, and without any additional noise parameters (i.e., using the SNR margin only) we proceed with a Lagrange relaxation based approach for the near-optimal solution of the original multi-carrier problem in (7.4).

Algorithm 8 DSM Scheme for the Joint Optimization of AN and Transmit Power

---

1: **while** Master Problem (Maximization in (7.8)) not Solved **do**
2:   Generate a New Set of Dual Variables $\boldsymbol{\lambda}$ and $\boldsymbol{\nu}$ (e.g.,
     By the LP based Column-Generation Scheme in Section 4.1.1)
3:   **for** All Subcarriers $c \in \mathcal{C}$ **do**
4:     **while** Optimal Bit- and Power-Allocation Not Found **do**
5:       Follow an optimal, discrete search method (e.g., the BnB Search in Section 3.1.2.1) to Obtain Another $\mathbf{r}^c$
6:       Evaluate $\mathbf{p}^c(\mathbf{r}^c)$ and the Objective in (7.8a) by Solving the LP in (7.6) with weights $\mathbf{w} = (\hat{\mathbf{w}} + \boldsymbol{\nu})$
7:     **end while**
8:     *Optional* (for the specific multiplier search scheme in Section 4.1.1, cf. Line 2 above): Recover a solution to the problem in (7.4) by the Heuristic in Algorithm 2
9:   **end for**
10: **end while**

7.2 Performance Bound Computation for AN-enabled Networks

In this section we approach the problem in (7.4) by solving its partial Lagrange-dual problem [12]. As in Section 1.2.4, the Lagrange relaxation is motivated by the typically large number of subcarriers $|\mathcal{C}|$ which after the relaxation become independent in terms of the power-allocation. The dual problem is, similarly as in (1.10) for the case without AN, defined as

$$\underset{\boldsymbol{\nu},\boldsymbol{\lambda}}{\text{maximize}} \; \underset{\mathbf{p}^c \succeq \mathbf{0}, \mathbf{a}^c \succeq \mathbf{0}, c \in \mathcal{C}}{\text{minimize}} \; \sum_{c \in \mathcal{C}} f^c(\mathbf{p}^c, \mathbf{a}^c, (\hat{\mathbf{w}} + \boldsymbol{\nu}), (\breve{\mathbf{w}} + \boldsymbol{\lambda})) + \sum_{u \in \mathcal{U}} \nu_u \hat{P}_u - \lambda_u R_u \quad \text{(7.8a)}$$

subject to   Per-subcarrier constraints in (7.4c) – (7.4d),   (7.8b)

where $\boldsymbol{\nu}, \boldsymbol{\lambda} \in \mathbb{R}^U$ are the Lagrange multipliers associated with the relaxed sum-power and sum-rate constraints in (7.4b), respectively.

7.2.1 An Optimal Algorithm for the Dual Problem in (7.8)

Our scheme for optimally solving the problem in (7.8) is summarized in Algorithm 8. The minimization part in (7.8a) is solved by a discrete search over the bit-allocation $\mathbf{r}^c, c \in \mathcal{C}$, instead of an optimization over the (continuous) spectral power-allocation variables with discrete-rate constraints as in (7.4c), as suggested by (7.8a). For this purpose we use the problem-specific and optimal DFB search scheme specified in Section 3.1.2.1, where we note that any other discrete search scheme such as that

in [163] can in principle be used instead. Furthermore, such a discrete search involves the evaluation of the powers $\mathbf{p}^c(\mathbf{r}^c)$ by solving the LP in (7.6) in order to evaluate feasibility in (7.8b) and the objective in (7.8a). On top of the discrete search for the bit-allocation comes an iterative scheme targeting the maximization in (7.8a) (the "master problem"), where we use the LP based column-generation scheme specified in Section 4.1.1. Again we note that other algorithms for non-differentiable optimization problems [12, Ch. 6] can in principle be applied instead, such as an exhaustive search [33], a subgradient search [163], or the ellipsoid method suggested in [198]. Altogether, we see that the optimization of AN jointly with the spectral transmit-power allocation neatly integrates into previous DSM approaches, notably also into low-complexity heuristics as in Section 4.2, cf. also Section 7.2.3 for an extension of multi-user bit-loading [103]. While in this section we assumed a fixed margin $\gamma_u, u \in \mathcal{U}$, the spectral AN allocation and the SNR margin are coupled and are therefore jointly optimized in Section 7.2.3. The optimal scheme of this section is only applicable to problems with a few users. However, simplifications as the introduction of "virtual disturber lines" [29, 104] reduce the number of lines which are jointly optimized and therefore make such an optimal scheme also practically relevant. Similarly, in Section 7.2.5 we simplify the multi-user optimization problem by assuming an equal AN and transmit-power allocation for all collocated lines.

7.2.2 Simulation Results for Multiple Users and Fixed Margins

We evaluate the performance of our joint AN and power optimization scheme in Algorithm 8 in a 3-user downstream ADSL2 scenario[2] with loop-lengths of 800, 1100, and 1400 m, respectively. By weak-duality [12] we find that the suboptimality of the primal solutions for the original problem in (7.4) obtained by the application of the heuristic in Algorithm 2 subsequent to our dual-optimal algorithm (cf. Line 8 in Algorithm 8) is below 10^{-4} % in all the simulation results shown in this section. This means that the shown solutions, which apply to the original problem in (7.4), are provably near-optimal. Figure 7.2(a) shows the optimal mean rate among all users over the SNR margin γ, which is set equal for all users $u \in \mathcal{U}$ and increased in steps of 0.2 dB. We see that the performance under AN increases up to a certain margin value (at around $\gamma = 12.6$ dB), beyond which it decreases again. This behavior can be intuitively explained by the constraints in (7.6c) which limit the SNR variation under crosstalk and therefore, for small values of γ, the bit-allocation of heavily interfering lines. Differently, under VN we have constant performance up to $\gamma = 12.2$ dB which is explained by the fact that VN can fully replace the role of the SNR margin. Note

---

[2]The parameters for ADSL2 follow the standard in [86, Annex B], using non-overlapping spectra with ISDN, $\Gamma = 6.8$ dB, $\hat{\theta} = 15$, $\theta = 1$, and $N_u^c = -120$ dBm/Hz, $\forall c \in \mathcal{C}, u \in \mathcal{U}$. Weights for rate maximization are set to $\hat{w}_u = 0$, $\breve{w}_u = 1$, $u \in \mathcal{U}$, and to $\hat{w}_u = 1$, $\breve{w}_u = 0$, $u \in \mathcal{U}$, for sum-power minimization. The parameters for ADSL2+ follow the standard in [87, Annex A], using non-overlapping spectra with ISDN and other settings as for ADSL2.

that for better comparability to AN we always plot the transmitter-referred VN v_u^c/H_{uu}^c. Considering the third curve where the lines are stabilized by an adequate bit-allocation and the SNR margin only, we see that the optimal margin is higher than under AN (at around $\gamma = 14.8\,\mathrm{dB}$). The gain in sum-rate by AN compared to the optimal margin setting without extra noise is approximately $6.8\,\%$ (cf. the interval "(A)" in Figure 7.2(a)), while that of VN compared to AN is less than $0.4\,\%$. Note that this is a simplified evaluation of the gain by AN as we set the SNR margin equal for all users. We will perform a heuristic setting of the margin for each user separately in Section 7.2.3. Figure 7.2(b) further shows how for large values of γ the total used AN (as well as transmitter-referred VN) decreases. This is intuitive as for large values of the margin γ the initial bit-loading constraint in (7.6b) becomes more and more active while the SNR variation constraint in (7.6c) becomes more and more inactive and AN hence meaningless.

We repeat the simulation under the same simulation setup for ADSL2+ which uses approximately double the spectral bandwidth compared to ADSL2, cf. Figure 7.3. The best found SNR margin values are now $9.2\,\mathrm{dB}$ and $6\,\mathrm{dB}$ under no extra noise ("margin only") and AN, respectively, cf. Figure 7.3. This means that the optimal margin values are now smaller than for ADSL2, while we find that the AN and VN sum-power values slightly increased compared to the results for ADSL2 in Figure 7.2(b) (result not shown). This can be explained by the higher bandwidth used in ADSL2+ compared to ADSL2 and the frequency selectivity of the channel, which results in a more frequency selective crosstalk noise in ADSL2+. The gain at the optimal margin value γ by AN compared to the case with no extra noise is now $7.5\,\%$ (cf. the interval "(A)" in Figure 7.3), and that by VN compared to AN $4\,\%$, cf. Figure 7.3. Concluding, the total power spent for AN is negligible compared to the total transmit power.

Next we exemplarily have a look at the spectral shape of the optimized AN at our solution in Figure 7.3 for the margin $\gamma_u = 1\,\mathrm{dB}$, $\forall u \in \mathcal{U}$, in ADSL2+, cf. Figure 7.4 for an illustration. In [69] it is recommended that the (received) AN power levels should follow the spectral shape of the crosstalk noise (i.e., the optimal VN setting), and set at (or slightly below) the crosstalk power levels, as also described in [167]. However, from Figure 7.4 we see that this simple rule does not hold for every setting of the SNR margin. In addition, our simulations will show that the optimal margin value can vary widely for different loop lengths, cf. Section 7.2.4. Concluding, the power spectrum of the AN clearly needs to approximately follow that of the crosstalk noise in order to enforce the stability constraint in (7.6c). However, the precise/optimal relation between the AN and the crosstalk noise spectrum depends on the specified SNR margin.
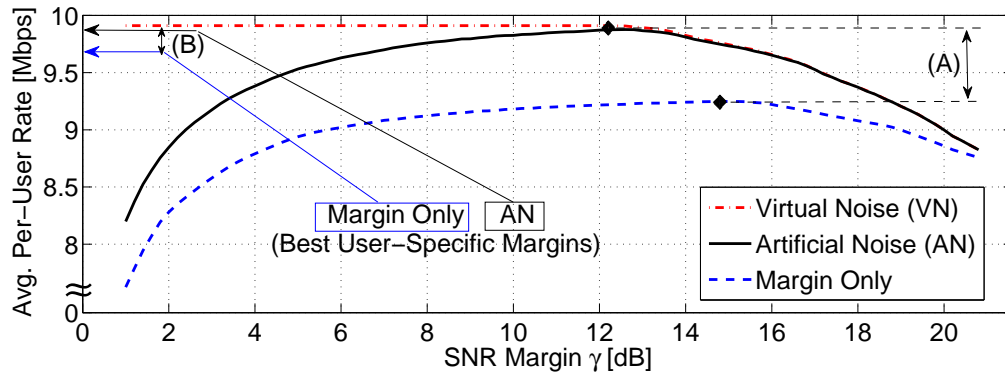
(a)
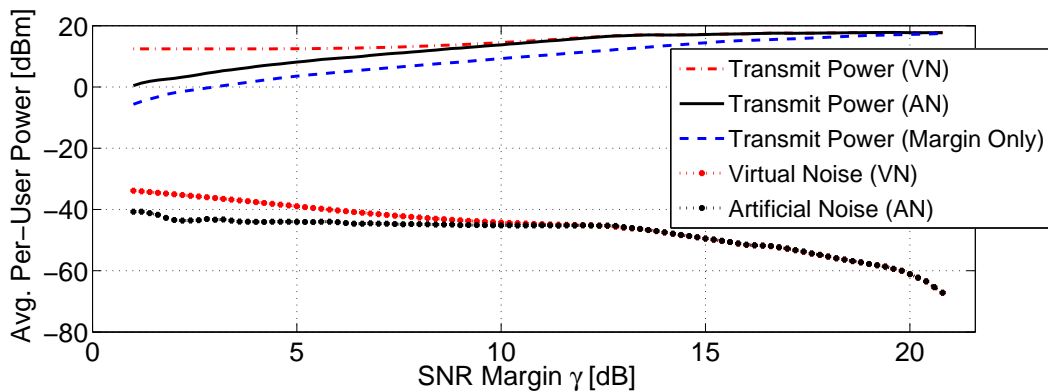
(b)

Figure 7.2: Dependency of the mean rate (a) and transmit powers (b) of three stabilization techniques on the set SNR margin $\gamma_u = \gamma, \forall u \in \mathcal{U}$ in ADSL2.

7.2.3 Heuristic for the Joint Optimization of Spectral Power-Allocation, Margin and Artificial Noise

Motivated by the dependency of the optimal spectral AN and transmit-power allocation on the SNR margin we propose a heuristic for jointly optimizing all three variable sets. However, the proposed scheme can also be applied to efficiently search for the single best margin for all users, as found in Section 7.2.2 by an exhaustive search. In order to obtain a low-complexity scheme we embed a multi-user bit-search technique in a search for the margins γ_u, performed sequentially one user after the other. A more explicit summary is given in Appendix C.3. We repeat the simulation setup of Section 7.2.2, this time optimizing the SNR margins on a per-user basis using Algorithm 15, initialized at $\gamma_u = 10\,\mathrm{dB}, \forall u \in \mathcal{U}$. Beginning with ADSL2 (cf. the

Figure 7.3: Dependency of the mean rate under three stabilization techniques on the set SNR margin $\gamma_u = \gamma, \forall u \in \mathcal{U}$ in ADSL2+.
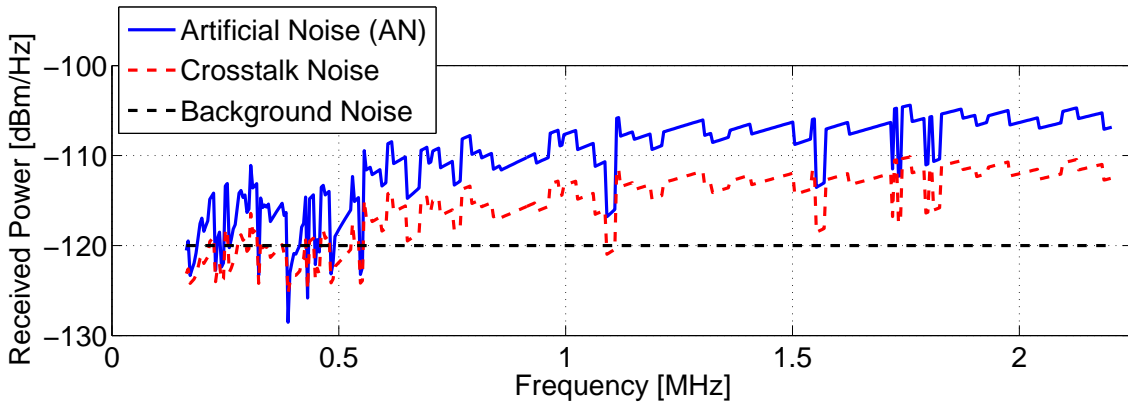
Figure 7.4: Relation between the crosstalk and the optimized artificial noise received on the shortest (800 m long) line for a margin of $\gamma_u = 1$ dB, $\forall u \in \mathcal{U}$, in ADSL2+.

results for equal margins among users in Figure 7.2) we find that the gain by AN compared to no extra noise ("margin only") is now less than 1.8 %, cf. the interval "(B)" in Figure 7.2(a). The little rate gains for AN by per-user margin optimization are explicable by the the fact that VN upper-bounds the AN performance for any margin setting under AN, and the performance under AN and a single margin was already close to that of VN in Figure 7.2(a). The results for user-specific margin optimization in Figure 7.3 show again higher improvements for the margin-only scheme, making the difference to the scheme with AN shrink to less than 3.8 %, cf. the interval "(B)" in Figure 7.3. Summarizing this section, AN seems to be able to partially compensate for the performance loss incurred by setting the SNR margin values of all users equal. However, also the single-margin value needs to be optimized under AN based on the actual network topology, a point we investigate further in the next section.
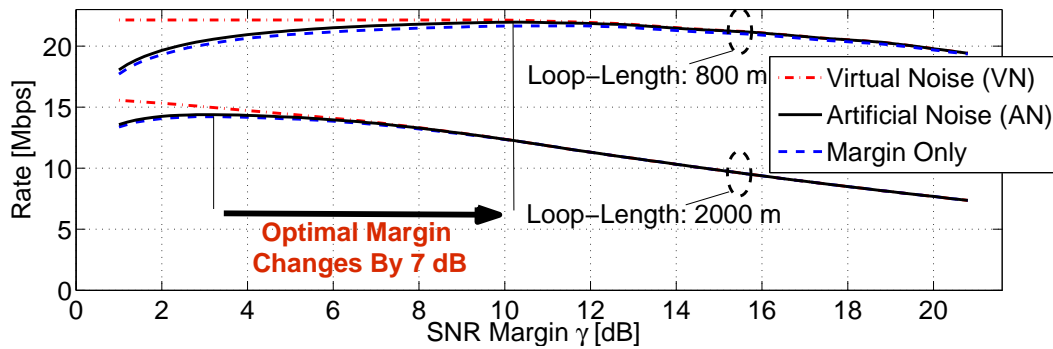
Figure 7.5: Dependency of the mean rate on the single SNR margin $\gamma_u = \gamma, \forall u \in \mathcal{U}$, for three stabilization techniques in two collocated scenarios.

7.2.4 Margin Optimization in Collocated Scenarios

Summarizing our simulation results so far, they indicate that a near-optimal (with reference to the VN scheme) stable rate is, in networks with limited distribution of users, already achievable by an adequate setting of a single network-wide SNR margin. In this section we investigate scenarios with 3 collocated users and loop-lengths of 800 m and 2000 m, respectively, and apply again the near-optimal algorithm of Section 7.2.1 for a varying SNR margin $\gamma_u = \gamma, \forall u \in \mathcal{U}$. We find that the per-user margin optimization heuristic of Section 7.2.3 does not significantly improve the sum-rate compared to the optimal single margin γ in this collocated setup, and omit these results for this reason. However, as seen in Figure 7.5 there is a large gap between the optimal single-margin settings for the two loop-lengths, supporting our point that the optimization of AN and the SNR margin need to be done jointly.

7.2.5 Simplification by Assuming an Equal Power-Allocation for Collocated Lines

In Section 7.2.1 we mentioned the simplification of the optimal solution of the Lagrange-dual multi-user DSM problem in (7.8) by considering "virtual lines", which represent several other, real lines in the system. Another simplification we may think of is to enforce an identical power-allocation (in terms of AN/VN and transmit-power/bit-allocation) for all collocated lines. This leads to small modifications of the subproblems in (7.6) or (7.7) in the sense that the crosstalk from each optimized user is multiplied by the number of lines collocated with the crosstalker, and additionally we need to consider the crosstalk from the lines that are collocated with the victim line. The latter crosstalk is determined by the power-allocation of the victim line, and can hence be interpreted as a "self-noise". As the collocated lines share the same solution, the optimization of a single line enforces the power and sum-rate constraints in (7.4b) for all collocated lines, assuming that the target-rates, sum-power
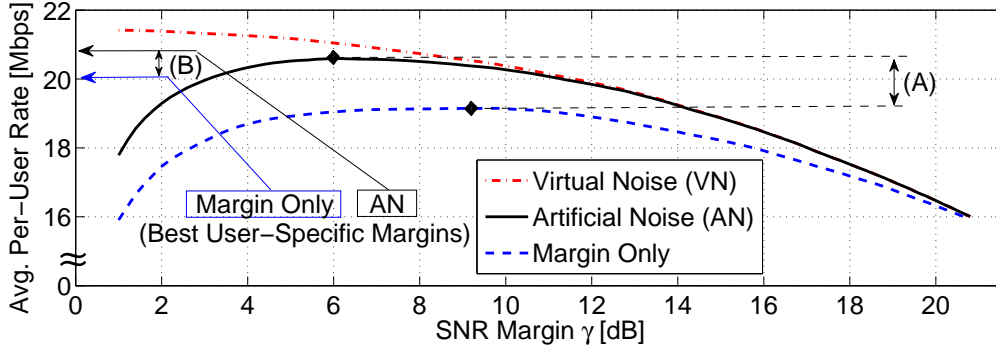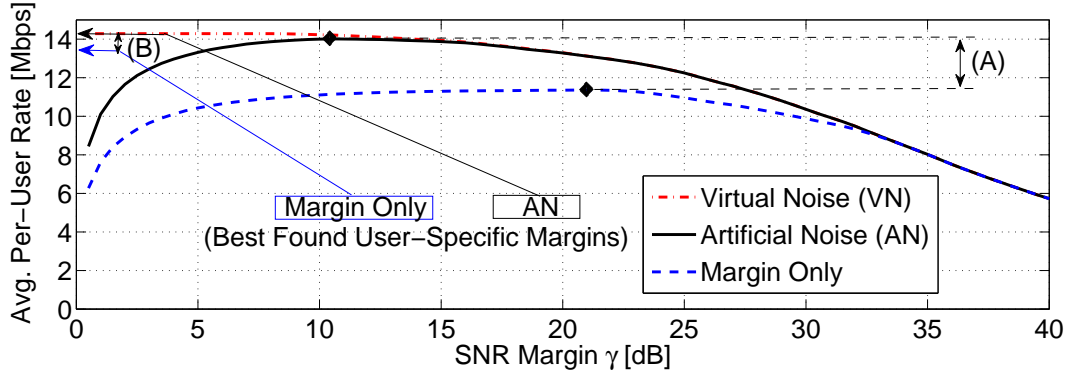
Figure 7.6: Dependency of the mean rate under three stabilization techniques on the set SNR margin $\gamma_u = \gamma, \forall u \in \mathcal{U}$ in an ADSL2+ network with 33 lines.

budgets, channel, DSL technology, etc., are equal (which we do here). The main benefit from this simplification is that it allows us to make a performance evaluation under optimal DSM in an environment with stronger crosstalk. In [30] a practical DSM implementation on standard-compliant DSL chipsets is demonstrated, and the application of a single (optimized) spectral mask to a group of users suggested to reduce the complexity of implementing DSM. Our simplification is closely related to this line grouping, where the crosstalk inside the group and into other groups is explicitly modeled.

We consider three collocated ADSL2+ user groups of 11 lines each (i.e., 33 lines in total), located, as above, at 800 m, 1100 m and 1400 m distance from the central office, respectively, and other simulation parameters as specified in Section 7.2.2.

The results shown in Figure 7.6 qualitatively resemble those in Figure 7.3. For the best single-margin setting we observe a rate and transmit power under AN which is similar to those under VN. Furthermore, again we see that the margin-only scheme profits most from a user-specific SNR margin, where the extra benefit by AN in terms of mean rate drops from more than 23 % to less than 5 %. Comparing Figures 7.6 and 7.3 we find that the higher crosstalk level in this 33-user example compared to the 3-user example in Section 7.2.2 results in a higher optimal SNR margin and a higher AN sum-power (result not shown) at the optimal margin. Note also that the extra benefit by VN compared to AN is below 2 % (or 0.3 Mbps). We repeated this simulation, this time with all 33-users being collocated at a distance of 2000 m from the deployment point (results not shown). While the optimal SNR margin for AN approximately halves, the benefit by VN compared to AN is now more than 11 % (or more than 1.1 Mbps). Regarding once more Figure 7.5 for 3 collocated, separately optimized ADSL2+ lines, we can draw a similar conclusion. There we find a benefit by VN of around 0.8 % (or 0.18 Mbps) for a loop-length of 800 m, and of more than 8 % (or 1.17 Mbps) in case of 2000 m. Summarizing, compared to the "ideal" frequency selective SNR margins (that is VN) the stable performance
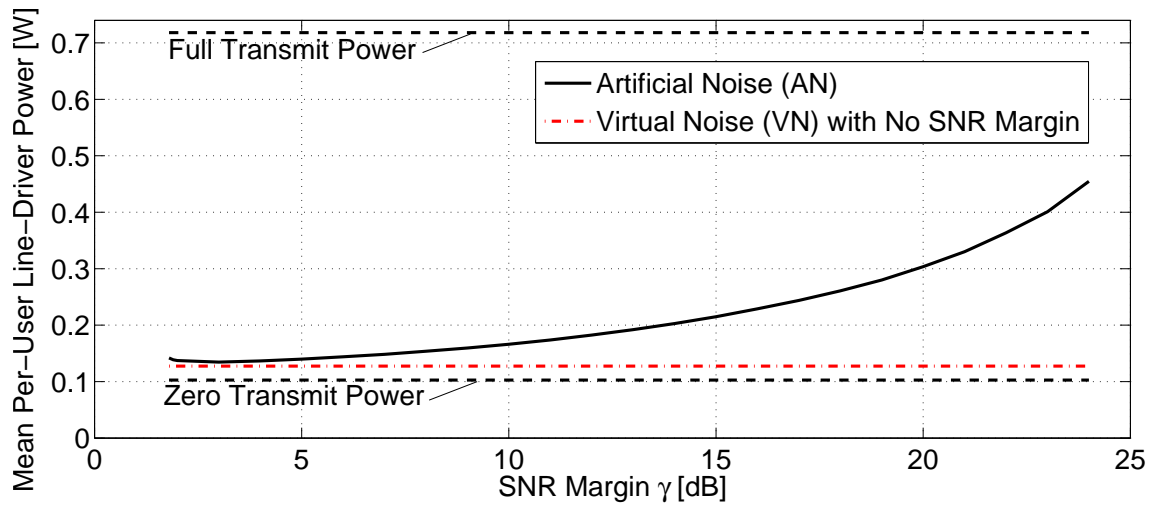
Figure 7.7: Dependency of the mean line-driver power consumption on the single SNR margin $\gamma_u = \gamma, \forall u \in \mathcal{U}$, under AN compared to the optimal power under VN in a 33-user ADSL2+ scenario.

suffers from the injection of AN especially at longer loop-lengths. This effect will be further analyzed in Section 7.3.2 by loop-reach simulations, where we will also give an intuitive explanation.

In Figure 7.7 we look at the mean line-driver power consumption[3] for supporting 80 % of the per-user rates achieved by a sum-rate maximization under VN. While the power under AN is seen to be fairly margin dependent, under the best shown margin setting (at $\gamma = 3\,\mathrm{dB}$) it is close (less than 6 % higher) to that under VN. Furthermore, the optimal margin value is far below that in the corresponding sum-rate maximization problem (cf. Figure 7.6), meaning that the optimal margin setting is not only channel and network topology, but also target-rate dependent. A word of caution is needed at this point, as the typically encountered low AN sum-power level (in comparison to the total transmit power) does *not* automatically imply that the AN technique is energy-efficient, as transmit power may also be wasted to compensate for the needed frequency-flat SNR margin. We will see an example supporting this claim in the next section where a power-mask based crosstalk noise is assumed in order to decouple and therefore simplify the original multi-user DSM problem in (7.4).

---

[3]The used model in this chapter is that of a class AB line-driver and uses the parameterization specified in [185].

7.3 Performance Analysis under Worst-Case Crosstalk

A commonly applied simplification in performance evaluation for multi-user DSL networks is to decouple the users by assuming the highest possible crosstalk noise t_u into line $u \in \mathcal{U}$ based on the spectral mask constraints[4] in (7.6d), given as

$$\sum_{i \in \mathcal{U} \setminus u} H_{ui}^c \left(p_i^c + a_i^c \right) \leq t_u^c = \sum_{i \in \mathcal{U} \setminus u} H_{ui}^c \hat{p}_i^c. \tag{7.9}$$

Replacing the crosstalk terms by these (constant) upper-bounds in the per-subcarrier problem in (7.6), we see that the users' AN and power-allocation become decoupled problems, and the constraint in (7.6c) can be simplified to

$$H_{uu}^c a_u^c \geq \tilde{a}_u^c \doteq \left(t_u^c / (\gamma_u - 1) - N_u^c \right). \tag{7.10}$$

As a decrease in the AN value a_u^c makes the constraints in (7.6b) and (7.6d) less restrictive for the power-allocation variables p_u^c and the goal is to minimize the total transmit power in (7.6a), we can give the optimal AN setting explicitly as

$$a_u^c = \left[\tilde{a}_u^c \right]_+ / H_{uu}^c, \tag{7.11}$$

where $[x]_+ = x$ for $x \geq 0$, and $[x]_+ = 0$ otherwise.

7.3.1 Optimal Algorithm for Stable Bit-Loading

We investigate the setting of the AN jointly with the transmit-power allocation based on the decoupling worst-case assumption in (7.9), which, as will be shown, can be solved efficiently and optimally through a (modified) greedy bit-loading procedure. Inserting (7.11) back into (7.6) and regarding the original problem in (7.4) we see that one recovers the modified single-user bit-loading problem

$$\underset{\tilde{r}_u^c \in \mathcal{B}, c \in \mathcal{C}}{\text{minimize}} \quad \sum_{c \in \mathcal{C}} \hat{w}_u \left(p_u^c (\tilde{r}_u^c) + \left[\tilde{a}_u^c \right]_+ / H_{uu}^c \right) - \breve{w}_u \tilde{r}_u^c \tag{7.12a}$$

$$\text{s.t.} \quad \sum_{c \in \mathcal{C}} \tilde{r}_u^c \geq R_u, \quad \sum_{c \in \mathcal{C}} p_u^c (\tilde{r}_u^c) \leq \hat{P}_u, \tag{7.12b}$$

$$p_u^c (\tilde{r}_u^c) \leq \hat{p}_u^c - \left[\tilde{a}_u^c \right]_+ / H_{uu}^c, \quad \forall c \in \mathcal{C}, \tag{7.12c}$$

$$\text{where} \quad p_u^c (\tilde{r}_u^c) \doteq (2^{\tilde{r}_u^c} - 1) \frac{\gamma_u}{H_{uu}^c} \Gamma \left(N_u^c + \left[\tilde{a}_u^c \right]_+ \right), \tag{7.12d}$$

which we write in terms of the variables $\tilde{r}_u^c$ instead of p_u^c to emphasize the relation to

---

[4]Note however that estimates of the worst-case noise encountered in real networks are more commonly based on long-term network observations [167].

Algorithm 9 Single-user AN and Bit-Allocation Scheme

1: **for** all users $u \in \mathcal{U}$ **do**
2:    **while** (e.g., exhaustive) search for γ_u **do**
3:       Compute the constant received AN $\left[\tilde{a}_u^c\right]_+$ as in (7.10)
4:       Solve the problem in (7.12) by greedy bit-loading [27]
5:    **end while**
6: **end for**

---

bit-loading algorithms. AN solely leads to a modified (but constant) additive objective term, mask, and background noise in (7.12a), (7.12c), and (7.12d), respectively. Hence, we find that greedy bit-loading [27, 178] (with sum-power objective) is optimal when applied to the problem in (7.12). We refer to [27] for the corresponding result under a sum-power objective, and Appendix A.2 for a proof which is applicable to the cost function in (7.12a). The intuition is that the weights in (7.12a) are equal for all subcarriers and hence do not alter the optimal decisions the classical single-user bit-loading algorithm [27] takes. This optimal greedy bit-loading procedure needs to be embedded in a (e.g., exhaustive) search loop for the SNR margin value γ_u, cf. Algorithm 9 for a generic description.

The same arguments can be applied to show that the single-user bit-loading problem of optimizing VN (under general margins $\gamma_u \geq 1$ and $v_u^c \geq 0$) is readily solvable by a (modified) greedy bit-loading algorithm, and the (receiver-referred) VN power-allocation explicitly given as $v_u^c \doteq \left[\frac{t_u^c}{\gamma_u} - \frac{\gamma_u-1}{\gamma_u}N_u^c\right]_+$, which for $\gamma_u > 1$ is equivalent to $(\gamma_u - 1)/\gamma_u\tilde{a}_u^c$. This implies that, assuming worst-case crosstalk as in (7.9) and an identical SNR margin, the VN spectral power-allocation is always below that of the AN. Assuming the (for VN) optimal selection $\gamma_u = 1$, the VN v_u^c equals the crosstalk noise t_u^c, as already remarked in Section 7.1.5. As in the case of AN, the bit-loading considers a background noise increased by the additional noise v_u^c, while the objective and power mask constraints are, differently to (7.12a) and (7.12c), not altered by the VN.

We proceed with simulation results where the proposed modified greedy bit-loading procedure is applied to a single user, showing the optimal rate and sum-power levels under the conservative crosstalk assumption in (7.9).

7.3.2 Evaluation of Stable Single-User Bit-Loading

We repeat the 3-user ADSL2+ example of Section 7.2.2, this time using the worst-case crosstalk assumption of Section 7.3 and Algorithm 8 applied to each user individually. The results are shown in Figure 7.8, cf. Figure 7.3 for the corresponding results under our near-optimal multi-user DSM scheme. First note that without the extra noise (AN or VN) and fixed crosstalk noise the feasibility of the stability constraint in (7.6c) solely depends on the set SNR margin. Hence, in Figure 7.8 we
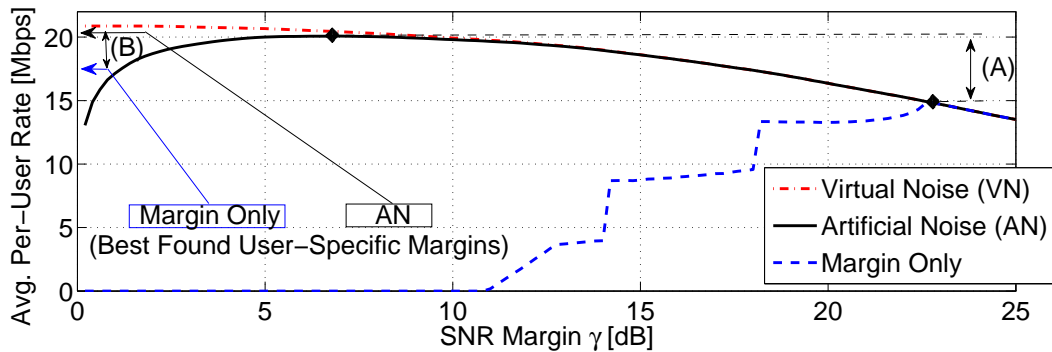
(a)
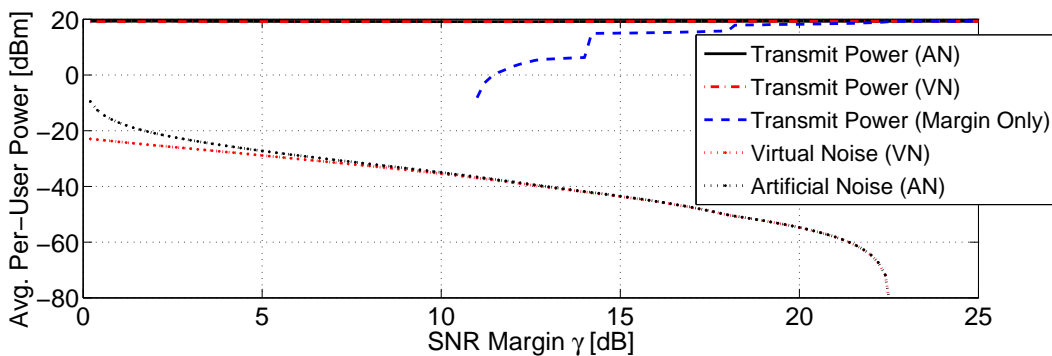
(b)

Figure 7.8: Average rates (a) and power levels (b) under three stabilization tech-
niques and a worst-case crosstalk assumption in a distributed 3-user
ADSL2+ scenario.

obtain no (stable) rate below a certain threshold margin. Considering the larger
crosstalk noise considered here it comes as no surprise that the achieved rates are
lower and the AN and VN power levels higher in Figure 7.8 compared to Figure 7.3.
In Figure 7.8(a) we additionally show the mean rate achieved when we are allowed
to select the SNR margin for each user individually and optimally (up to the selected
granularity of 0.2 dB). Similarly as in Section 7.2.3 we find that this additional free-
dom in setting the margin mostly improves the performance in the case when no
extra noise (AN or VN) is used. Note that due to the user-independence under the
worst-case assumption in (7.9) it comes at no additional effort to find the optimal
user-specific margins compared to finding the single optimal margin applied to all
lines.

The main advantage of the conservative simplification made in in this section is
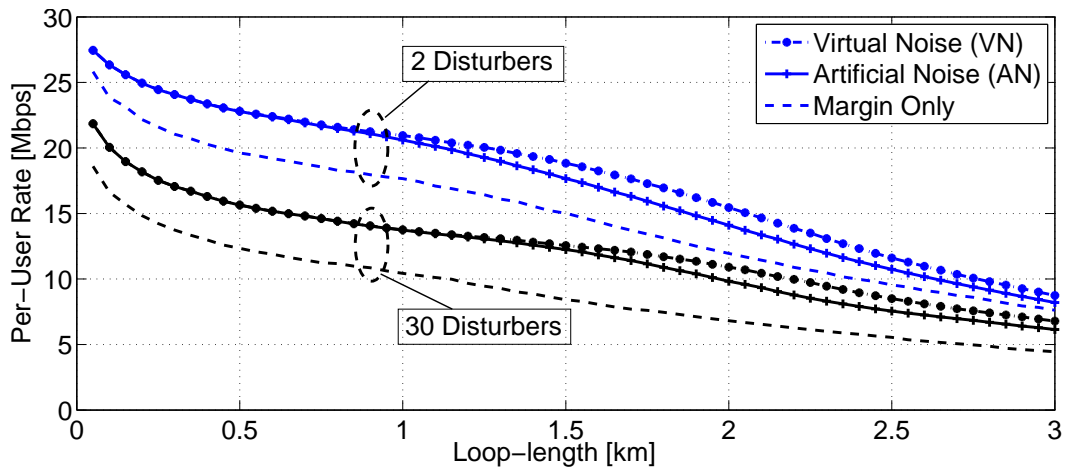
Figure 7.9: Rate under the three stabilization techniques in ADSL2+ for 2 and 30 collocated disturbers, respectively.

that it allows to do a performance evaluation in networks with a large number of users. Hence, we do now assume numerous disturbers collocated with an ADSL2+ system, and simulate Algorithm 8 with an exhaustive search over the margin with a granularity of $0.2\,$dB. The results in Figure 7.9 represent the rate performance for a certain loop-length and the best SNR margin selection in the above sense for 2 and 30 collocated disturbers, respectively.[5] Compared to the scenario without extra noise we find that AN provides a gain in rate of between $6.3\,\%$ and $18.9\,\%$ in case of 2 disturbers, between $13.4\,\%$ and $36.7\,\%$ in the case of 10 disturbers, and between $17.4\,\%$ and $48.2\,\%$ in case of 30 disturbers. Furthermore, AN gives a (worst-case crosstalk) rate performance similar to VN, at least for loop-lengths below 1 km. Above this length we find a maximum rate loss compared to VN of $8.9\,\%$, $11.3\,\%$, and $12.4\,\%$ for 2, 10, and 30 disturbers, respectively.

In order to further study the energy-efficiency of AN we repeat this loop-reach simulation, but show instead of the achieved rate the line-driver power consumed at $80\,\%$ of the maximum rate achieved by the VN technique. The results depicted in Figure 7.10 show a loss in terms of line-driver power consumption by AN compared to VN for long loops, similar as observed in terms of rates in Figure 7.9. This can be explained intuitively as for long loops the performance becomes less constrained by the crosstalk noise t_u^c but more constrained by the noise $H_{uu}^c a_u^c + N_u^c$, while VN does not suffer from the artificial noise term $H_{uu}^c a_u^c$. The maximum extra power cost by AN compared to VN is around $27\,\%$ (or $119\,$mW, respectively) for a loop-length of $2.25\,$km. Note however that the extra "power cost" by AN becomes less evident for lower target-rates (e.g., below $50\,\%$ of the maximum rates, results omitted) where the power consumption is lower as well. Figure 7.10 also shows the power consumption

---

[5]Results obtained for 5 and 10 disturbers are qualitatively similar (although shifted in terms of rates), and therefore omitted.
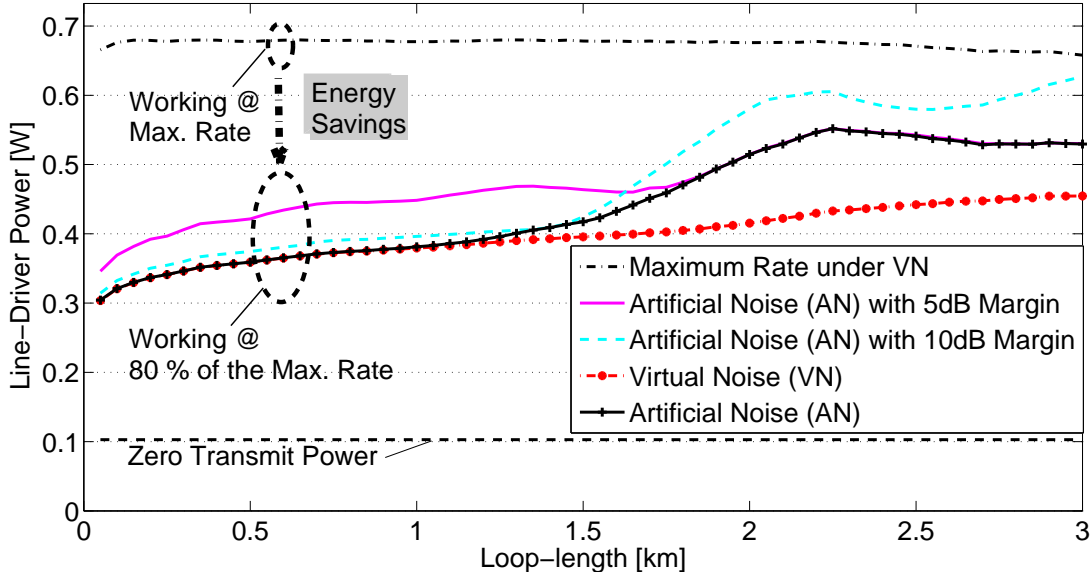
Figure 7.10: Line-driver power consumption of a single line under study, operating at 80 % of the maximum achievable stable rate, under the three studied stabilization techniques and 30 collocated disturbers in ADSL2+.

for two fixed values of the SNR margin. The comparison to the line-driver power consumption under AN and the best selected margin shows once more the gain by SNR margin optimization under varying loop-lengths.

7.3.3 Stable Sequential Initialization under Worst-Case Crosstalk

In Section 7.3.2 we evaluated the performance of a line seeing no crosstalk noise during its initialization, which is stabilized for the worst-case crosstalk noise in (7.9). Differently, in case a line u initializes when other lines are already active we need to consider the additional crosstalk noise $\tilde{t}_u^c, c \in \mathcal{C}$, into line $u \in \mathcal{U}$ that is caused by the active disturbers. Hence, the crosstalk noise present at initialization also needs to be added to the right-hand side of the stabilization criterion in (7.6c). This leads to a lower optimal AN level (cf. (7.11))

$$a_u^c \doteq \frac{1}{H_{uu}^c}\left[\left(\frac{t_u^c - \gamma_u \tilde{t}_u^c}{\gamma_u - 1} - N_u^c\right)\right]_+. \tag{7.13}$$

Using this definition the optimal modified bit-loading problem is given similarly as in (7.12), differing only in the higher background noise $(N_u^c + \tilde{t}_u^c)$ and the altered AN levels in (7.13).

The noise levels $\tilde{t}_u^c$ at initialization depend on many factors as mentioned in the
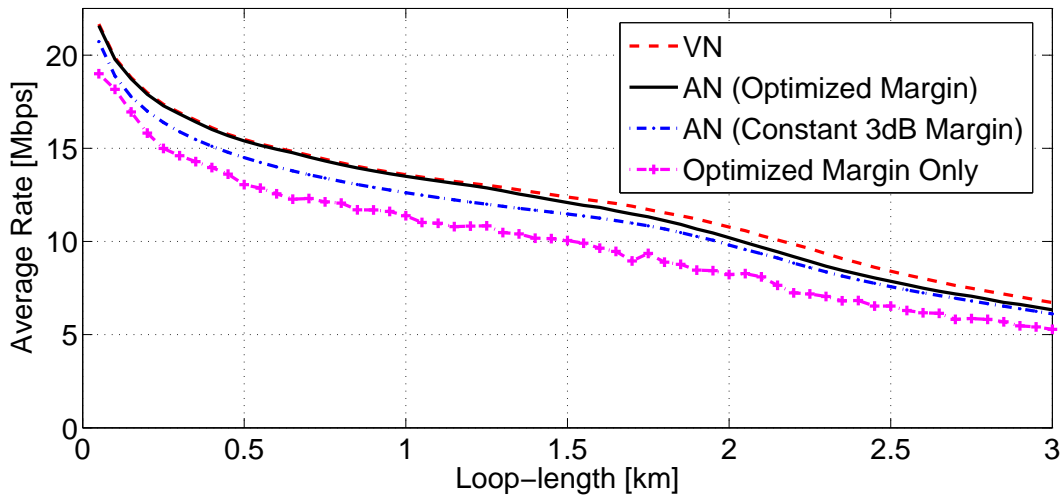
Figure 7.11: Average rate of 33 stabilized ADSL2+ lines under sequential initialization.

introduction to this chapter, including the users' line usage behavior. As a simple possible rule for performance simulation we assume that the lines initialize sequentially, one user after the other, as also assumed in [167]. In a collocated scenario this leads to a unique average rate per line, where we apply the modified bit-loading algorithm in total U times with sequentially updated received crosstalk levels $\tilde{t}_u^c$, and AN levels in (7.13). In Figure 7.11 we show the obtained simulation results for collocated 33-user scenarios with varying loop-lengths. The lowest and highest curves depict the performance without AN/VN and with VN, respectively. Especially at lower loop-lengths we find a noticeable gain by margin optimization under AN. More precisely, we find a maximum gain in average bit-rate in Figure 7.11 compared to the fixed margin setting at 3 dB suggested in [167] of over 7 %. In terms of reach the gain is even beyond 32 % up to a loop-length of 1 km.

7.3.4 Comparison of Performance Evaluation Techniques

In Figure 7.12 we compare the three proposed evaluation techniques in terms of loop-reach simulations: The DSM based AN setting for collocated scenarios of Section 7.2.5, the single-user optimization under worst-case crosstalk noise stabilization criterion in Section 7.3.1, and the sequential initialization scheme with modified bit-loading in Section 7.3.3. While the combination of AN and DSM gives rates that are partly (for longer loops) below those under those obtained by sequential initialization and worst-case crosstalk noise stabilization, they form an upper bound for the single-user worst-case stabilization scheme. The explanation is that while for both, DSM and single-user optimization, we consider a stabilization criterion for the case where no crosstalk noise is present at initialization, the DSM scheme results in an
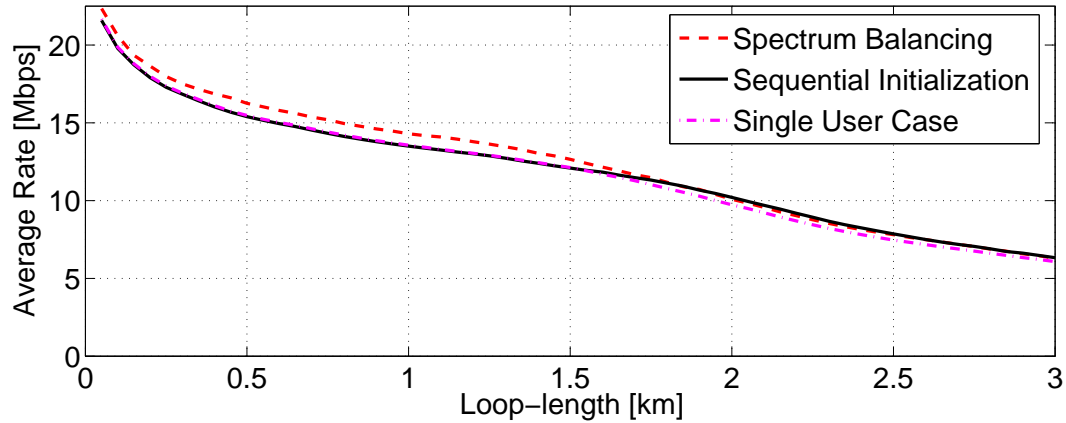
Figure 7.12: Average rate of 33 artificial noise enabled ADSL2+ lines under different performance evaluation techniques.

optimized crosstalk noise spectrum and hence higher performance. Differently, all but one line in the sequential evaluation scheme see a non-zero crosstalk noise power at initialization, and the stabilization criteria are hence less restrictive for most lines. Note however that the stabilization constraints in (7.6c) are easily adaptable to this scenario and DSM is hence also usable to bound the performance under sequential initialization.

Chapter 8 Conclusions

8.1 Summary of the Thesis

We considered the use of dynamic spectrum management (DSM) for the reduction of the power consumption in interference-limited multi-carrier digital subscriber lines (DSL). In the first part of the thesis we motivated our approach, which is based on the constrained optimization of the modems' transmit power. The line-driver is responsible for a major part of the transceiver's total power budget. However, most current low-complexity DSM algorithms are designed for the minimization of the transmit power. We presented empirical evidence which shows the efficiency of transmit-power minimizing DSM in reducing the line-driver power consumption. The DSM problem has an exponential complexity in the number of subcarriers. Lagrange-dual relaxation has previously been proposed as a means to approximate the original problem with one that has a linear complexity in the number of subcarriers. However, this approximation theoretically incurs a gap in the optimal objective value compared to the original DSM formulation. We have shown that this gap is indeed non-zero and dependent on the target-rate, but also found that it is negligible in typical DSM problems of practical size.

In the second part of the thesis we focussed on novel DSM algorithms. We have shown that the decomposed combinatorial per-subcarrier power control problems have a complexity which is polynomial in the number of DSL lines. However, as common search methods are still too complex for moderately sized DSL networks, we aimed at optimal search schemes with even lower complexity. This was achieved by exploiting the typically low crosstalk couplings in the search algorithm. Targeting DSM problems in even larger networks we focussed on the design of various per-subcarrier bit-search heuristics. Differing from the commonly applied Lagrange relaxation, we demonstrated a nonlinear Dantzig-Wolfe decomposition based DSM framework which is able to coordinate the possibly suboptimal solutions obtained by these heuristics in a robust fashion. We also proposed a novel coverage optimization problem which we approached by a linear as well as a semidefinite problem relaxation. Its solution is based on our DSM framework, which also formed the basis for the solution of the DSM problems below.

The third part of the thesis was devoted to stabilization techniques, including the proposal of a novel concept of crosstalk margins which relies on the solution of efficiently solvable convex per-subcarrier problem formulations. Furthermore, we analyzed the performance of DSL networks which are enabled to add artificial

noise at the transmitter. This artificial noise effectively stabilizes the network performance and therefore enables the energy savings promised by low-power modes in DSL.

8.2 Main Contributions of the Thesis

Key findings and conclusions are listed in the following.

Energy-Efficient DSM: We have empirically shown that optimizing the transmit power in the design and application of energy-efficient DSM algorithms instead of optimizing under the specific DSL line-driver incurs a negligible performance loss. The optimization towards transmit power leads to a problem which is decomposable into per-subcarrier problems and amenable to previously proposed low-complexity DSM algorithms.

Enhancing DSM by power-allocation in time has previously been considered for interference-limited DSL networks. We connected the possible energy savings by this enhancement to the duality-gap of the DSM problem, which we empirically found to be negligible for practical DSL systems. Hence, time-sharing in DSL is merely an algorithmic detour as demonstrated in this thesis, and yields little additional energy savings on top of that obtainable by DSM.

Energy-efficient DSM has been analyzed in terms of the line-driver power consumption under real-world channel data and in numerous randomly sampled near-far scenarios. We found that the line-driver power can on average be lowered by up to 70 % while still providing 80% of the maximum bit-rates.

Optimal Discrete-Rate Power Control: Contrary to previous belief a combinatorial single-carrier bit- and power-allocation problem subject to interference was seen to have a polynomial complexity in the number of users. Optimal search algorithms were proposed for its solution. A problem-specific search-space reduction method was proposed which demonstrates that weak crosstalk couplings can be algorithmically exploited to lower the complexity. By our optimal methods we have proved the near-optimality of previously proposed low-complexity, greedy DSM schemes in specific sum-rate maximization problem instances under real-world channel data.

Dantzig-Wolfe Decomposition: A novel framework for DSM based on a non-linear Dantzig-Wolfe decomposition was introduced, and the relation to the commonly applied Lagrange-dual relaxation demonstrated. Due to its robustness with respect to non-optimal subcarrier solutions it allows to perform a heuristic, discrete

bit-search and thereby to obtain discrete bit-allocations as a result. A combination heuristic was proposed which allows to return a feasible solution to the original problem, which is not available for classical Lagrange-dual relaxation based DSM schemes.

Our framework uses a storage of possible power-allocations. We demonstrated that this storage allows to fast and approximately resolve the DSM problem for various target-rates, which makes it an attractive basis for adaptive DSM solutions. Similarly the DSM problem is fast, but only approximately, re-solvable for different objective functions. In summary, this framework provides the foundation for the optimization of various DSM objectives, and using optimal as well as suboptimal power control schemes.

Coverage Optimization: The maximization of service coverage was formally introduced as a novel DSM objective and two heuristic solution approaches were compared. Furthermore, a dual-optimal algorithm was proposed which allows for the exact quantification of the suboptimality of these heuristics in small networks. Performance gains by the heuristics compared to sum-rate maximization in terms of average coverage and power consumption were demonstrated in large networks.

Stabilized DSL Networks: Models of crosstalk margins were studied and integrated in energy-efficient DSM algorithms. Results show that robust DSM solutions still allow to harvest the majority of the possible power savings by applying DSM while providing stable rates under worst-case crosstalk noise.

The evaluation of DSL lines by jointly performing DSM and the optimization of the artificial noise (AN) was proposed, giving novel means to evaluate a bound on the performance of AN-enabled networks. An analytical solution for the optimal single-user AN setting was derived. This enables the proposed low-complexity joint optimization with the SNR margin, which enhances the previously proposed way of operating AN-enabled networks. Our results show that a significant power reduction is achievable by low-power modes stabilized by AN, despite the additional power consumption caused by AN.

8.3 Suggestions for Future Research

- Future research on energy-efficient DSL protocols may be directed towards quantifying the theoretical gains by multi-user coordination of traffic-aware rate scheduling decisions, which are currently decentralized by the use of autonomous low-power modes, and which may strongly depend on the DSM coordination level.

- As we have seen in this thesis, the discrete-rate power control problem is too challenging in general to be approached by direct search methods. However,

we have shown by our search-space reduction technique in Section 3.2 that the special case of a low crosstalk coupling can be exploited algorithmically to obtain optimal solutions at low complexity. Another special DSM problem in interference-limited DSL networks has been studied in [162], where the per-user sum-power constraints in (1.8c) were relaxed to a total sum-power constraint for all users. The corresponding relaxed DSM problem under continuous power-loading was shown to be optimally solvable by convex optimization, given that certain conditions on the channel data hold. The identification of further special cases and algorithmic tools to lower the complexity of the optimal DSM problem in interference-limited networks, as well as the analytical quantification of the suboptimality in search heuristics, are two important directions for future research.

- In Section 4.2.7 we proposed a heuristic which exploits the similarity of the subchannels in DMT-based DSL systems. Similarly, in [66] a low-complexity DSM scheme is proposed where the solutions on a subset of the subcarriers are interpolated to obtain the power-allocation on the remaining subcarriers, thereby reducing the number of optimization variables. In Section 7.2.5 we applied a near-optimal algorithm to an approximated DSM problem, where it is assumed that all collocated DSL subscribers have an identical power-allocation. In all these examples the similarity of bit- or power-allocations among subcarriers or DSL lines has been exploited in some way. However, a unified framework on how to exploit this similarity best is an open research problem.

- We saw in Section 3.1.2.1 that dual variable information stemming from the linear master problem in (4.1) can be used to reduce the search complexity for solving the decomposed per-subcarrier problems. Similarly, in [100] an OFDMA system with a single sum-power constraint for all users was studied and the search complexity of the per-subcarrier problems reduced by using the information about previously tested dual variables. More research is needed in deriving algorithms with closer connections between the dual-variable optimization and the per-subcarrier problems to lower the search complexity in optimal bit- and power-allocation.

- A fundamental modeling assumption in this work is that of synchronization among the controlled DSL lines. The effect of asynchronism was analyzed in [34] and modeled as crosstalk among adjacent frequency subcarriers, which destroys the validity of assuming orthogonality among subcarriers. The DSM approach in [34] consists of first solving a DSM problem assuming synchronism, followed by a heuristic gradient-based bit-loading process for the actual DSM problem with asynchronous crosstalk. It remains an open question to which extent DSM schemes based on Lagrange relaxation or nonlinear Dantzig-Wolfe

decomposition as proposed in Chapter 4 are efficiently modifiable to the case of asynchronous crosstalk.

- In our analysis of the coverage maximization problem we proposed two master problem formulations which are based on two different problem relaxations. The formulation has an impact on the collection of power-allocations obtained by our Dantzig-Wolfe decomposition based DSM framework. Furthermore, based on this collection a selection heuristic was applied to maximize the service coverage. Further research is necessary to identify which are the most suitable master problem formulations for this purpose. Another extension of our work is the consideration of an arbitrary set of service-rate definitions.

Appendix A Proofs

A.1 Proof of Theorem 1

Proof. The proof of the first statement follows from showing the equivalence between the dual linear program to (2.3) and the dual problem in (1.10), cf. [12, p. 652] for a similar proof. The continuous relaxation of Problem (2.3) is a standard linear program and strong duality is obtained under feasibility [14] which holds by assumption. Introducing Lagrange multipliers $\boldsymbol{\lambda}, \boldsymbol{\nu} \in \mathbb{R}_+^U$ and $\boldsymbol{\Phi} \in \mathbb{R}^C$ for constraints (2.3b), (2.3c) and (2.3d), respectively, we obtain the dual linear problem as

$$\underset{\boldsymbol{\lambda} \succeq \mathbf{0}, \boldsymbol{\nu} \succeq \mathbf{0}, \boldsymbol{\Phi}}{\text{maximize}} \quad \sum_{u \in \mathcal{U}} \left(\lambda_u R_u - \nu_u \hat{P}_u \right) + \sum_{c \in \mathcal{C}} \Phi_c \tag{A.1a}$$

$$\text{subject to} \quad (\boldsymbol{\lambda} + \check{\mathbf{w}})^T \mathbf{r}^c(\mathbf{p}^{c,i}) - (\boldsymbol{\nu} + \hat{\mathbf{w}})^T \mathbf{p}^{c,i} + \Phi_c \leq 0, \forall i \in \mathcal{I}^c, c \in \mathcal{C}, \tag{A.1b}$$

The otherwise unconstrained variables $\boldsymbol{\Phi}$ are upper-bounded by the constraints in (A.1b) and their sum is maximized in (A.1a). More precisely, each variable Φ_c has $\sum_{c \in \mathcal{C}} |\mathcal{I}^c|$ upper-bound constraints in (A.1b) and at optimum it attains the minimum of those. Consequently we can replace $\boldsymbol{\Phi}$ by these minimum upper-bounds and obtain

$$\underset{\boldsymbol{\lambda} \succeq \mathbf{0}, \boldsymbol{\nu} \succeq \mathbf{0}}{\text{maximize}} \quad \underset{\{i_c \in \mathcal{I}^c, c \in \mathcal{C}\}}{\min} \quad \left\{ \sum_{c \in \mathcal{C}} \left(\hat{\mathbf{w}}^T \mathbf{p}^{c,i_c} - \check{\mathbf{w}}^T \mathbf{r}^c(\mathbf{p}^{c,i_c}) \right) + \boldsymbol{\lambda}^T \left(\mathbf{R} - \sum_{c \in \mathcal{C}} \mathbf{r}^c(\mathbf{p}^{c,i_c}) \right) + \right.$$
$$\left. \boldsymbol{\nu}^T \left(\sum_{c \in \mathcal{C}} \mathbf{p}^{c,i_c} - \hat{\mathbf{P}} \right) \right\}. \tag{A.2}$$

As by definition $\mathbf{p}^{c,i_c} \in \mathcal{Q}^c, \forall i_c \in \mathcal{I}^c, \forall c \in \mathcal{C}$, we see the equivalence to the dual problem in (1.10) with the per-subcarrier objective defined in (1.9).

The second statement follows as (2.3) has $2U + C$ constraints and therefore a solution exists[1] that has at most this number of non-zero variables [14]. As $|\mathcal{I}_{\boldsymbol{\xi}}^c| \geq 1$, for all $c \in \mathcal{C}$, we can subtract $|\mathcal{C}| = C$ from the number of non-zero variables and obtain that the number of subcarriers $|\mathcal{C}_{\boldsymbol{\xi}}^+|$ where time-sharing occurs is at most $2U$, concluding the proof. A similar conclusion can be drawn by applying the Shapley-Folkman theorem [11, p. 374], cf. the geometric interpretations of (1.8) and (1.10) in [182]. $\square$

---

[1]Such a solution is referred to as "basic solution" in LP theory [14].

A.2 Proof of Theorem 2

Proof. We assume $U = 1$ and hence all vectors in this section are in $\mathbb{R}_+$. First we prove the optimality of a greedy discrete bit-loading (DBL) algorithm which, starting from $\mathbf{r}^c = 0, \forall c \in \mathcal{C}$, sequentially loads θ bits where they incur the least extra cost $f^c_{\mathbf{r}^c+\theta} = \hat{w}_u(\mathbf{p}^c(\mathbf{r}^c+\theta) - \mathbf{p}^c(\mathbf{r}^c)) - \check{w}_u\theta$. We define the set of possible bit-steps $\mathcal{E}$, $|\mathcal{E}| = \sum_{c \in \mathcal{C}} |\mathcal{Q}^c_{\mathbf{r}}|$, a function which assigns each element of $\mathcal{E}$ the associated cost $f^c_{\mathbf{r}^c}$, $\mathbf{r}^c$ being the rate after the bit-step, and the matroid $(\mathcal{E}, \mathcal{F})$, $\mathcal{F}$ being the set of all subsets of $\mathcal{E}$ [63]. Optimality of greedy DBL follows now from the optimality of greedily picking elements out of $\mathcal{E}$ [63, Sec. 7.5] and the monotonicity of $f^c_{\mathbf{r}^c}$ in $\mathbf{r}^c$. It follows that greedy DBL demands the minimum sum-power for a given target-rate. With this in mind we denote the optimum of (1.8) for neglected sum-power constraints (1.8c) by $P^*(\mathbf{R})$. We denote the maximal achievable rate by $\hat{\mathbf{R}} = \sum_{c \in \mathcal{C}} \max\{\mathbf{r}^c | \mathbf{r}^c \in \mathcal{Q}^c_{\mathbf{r}}\}$ and the minimum number of loaded bits after which all remaining possible bit-steps have a positive cost by $\check{\mathbf{R}} \in \mathbb{R}_+$, i.e., $\check{\mathbf{R}} = \min_{\mathbf{R}}\{\mathbf{R} | \exists \mathbf{r}^c \in \mathcal{Q}^c_{\mathbf{r}}, c \in \mathcal{C}, \sum_{c \in \mathcal{C}} \mathbf{r}^c \leq \mathbf{R}, f^c_{\mathbf{r}^c+\theta} > 0, \forall c \in \mathcal{C}\}$. By optimality we have that $P^*(\mathbf{R})$ is constant for $\mathbf{R} \leq \min\{\check{\mathbf{R}}, \hat{\mathbf{R}}\}$. We will therefore focus on the case $\check{\mathbf{R}} < \hat{\mathbf{R}}$, where $P^*(\mathbf{R})$ is strictly monotonously increasing in $\mathbf{R}$ for $\mathbf{R} > \check{\mathbf{R}}$ due to optimality of greedy DBL. Convexity of $P^*(\mathbf{R})$ over $\mathbf{0} \leq \mathbf{R} \leq \hat{\mathbf{R}}$ follows from

$$\alpha P^*(\mathbf{R} - k_\alpha\theta) + \beta P^*(\mathbf{R} + k_\beta\theta) \tag{A.3a}$$

$$\geq \alpha(P^*(\mathbf{R}) - k_\alpha\epsilon_\alpha) + \beta(P^*(\mathbf{R}) + k_\beta\epsilon_\beta) \tag{A.3b}$$

$$= P^*(\mathbf{R}) - \alpha k_\alpha\epsilon_\alpha + \beta k_\beta\epsilon_\beta \geq P^*(\mathbf{R}), \tag{A.3c}$$

where in (A.3a) we form a convex combination of any two target-rates in the given interval with lower and higher rate than $\mathbf{R}$ (i.e., $k_\alpha, k_\beta \geq 0$), respectively. In (A.3b) we use the optimality of greedy DBL, defining $\epsilon_\alpha = P^*(\mathbf{R}) - P^*(\mathbf{R} - \theta)$ and $\epsilon_\beta = P^*(\mathbf{R} + \theta) - P^*(\mathbf{R})$. In (A.3c) we use $\alpha + \beta = 1$ and $\alpha k_\alpha = \beta k_\beta$ as we are interested in the convex combination at rate $\mathbf{R}$, and again the optimality of greedy DBL implying $\epsilon_\alpha \leq \epsilon_\beta$. Next we regard any convex combination of feasible solutions to our primal problem in (1.8), which are represented by the 3-dimensional set in (2.5). By Carathéodory's theorem [12, Prop. B.6], for $U = 1$ any point in the convex hull of $\check{\mathcal{Q}}$ can be represented by the convex combination of at most four points in $\check{\mathcal{Q}}$. We define target-rates $\mathbf{R} \in \{\tilde{\mathbf{R}} | \tilde{\mathbf{R}} = k\theta, k \in \mathcal{Z}_+\}$ and pick any such combination with sum-power values $\hat{\mathbf{P}}_\alpha, \hat{\mathbf{P}}_\beta, \hat{\mathbf{P}}_\gamma, \hat{\mathbf{P}}_\delta$, with achieved sum-rates $\mathbf{R}_\alpha, \mathbf{R}_\beta, \mathbf{R}_\gamma, \mathbf{R}_\delta$, and with a set of weights $\alpha, \beta, \gamma, \delta \geq 0$, such that $\alpha \mathbf{R}_\alpha + \beta \mathbf{R}_\beta + \gamma \mathbf{R}_\gamma + \delta \mathbf{R}_\delta = \mathbf{R}$ and $\alpha \hat{\mathbf{P}}_\alpha + \beta \hat{\mathbf{P}}_\beta + \gamma \hat{\mathbf{P}}_\gamma + \delta \hat{\mathbf{P}}_\delta \leq \hat{\mathbf{P}}$. Referring to (1.8a) and (1.9) we obtain

$$\hat{\mathbf{w}}^T \left(\alpha\hat{\mathbf{P}}_\alpha + \beta\hat{\mathbf{P}}_\beta + \gamma\hat{\mathbf{P}}_\gamma + \delta\hat{\mathbf{P}}_\delta\right) - \check{\mathbf{w}}^T\mathbf{R}$$

$$\geq \alpha P^*(\mathbf{R}_\alpha) + \beta P^*(\mathbf{R}_\beta) + \gamma P^*(\mathbf{R}_\gamma) + \delta P^*(\mathbf{R}_\delta) \tag{A.4a}$$

$$\geq P^*(\alpha\mathbf{R}_\alpha + \beta\mathbf{R}_\beta + \gamma\mathbf{R}_\gamma + \delta\mathbf{R}_\delta) = \hat{\mathbf{w}}^T\hat{\mathbf{P}}^{\text{DBL}} - \check{\mathbf{w}}^T\mathbf{R}, \tag{A.4b}$$

where the second inequality follows from the convexity of $P^*(\mathbf{R})$ and where $\hat{\mathbf{P}}^{\mathrm{DBL}}$ is the minimal sum-power corresponding to a feasible solution for (1.8), obtainable by greedy DBL as explained above. Hence, if there exists such a convex combination which meets the sum-power and sum-rate constraints, we can compute a feasible solution to the primal problem in (1.8) by greedy DBL which has a lower or equal objective. The reverse holds as a primal feasible solution is in $\breve{\mathcal{Q}}$. The proof follows from tight duality between the optimization over convex combinations meeting the sum-power and sum-rate constraints as in the time-sharing problem in (2.3), and the dual problem in (1.10), cf. Theorem 1. $\qquad\square$

A.3 Proof of Theorem 3

Proof. We will analyze the number of feasible discrete power-allocations $|\mathcal{Q}|$ per-subcarrier in an interference channel. Note once more that subcarrier indices are omitted in this section for ease of notation. We show that there is a bound $\hat{U}$ on the maximum number of users which can jointly transmit at the lowest positive rate θ. This will be seen to limit the number of "types" of rate allocations by a scenario dependent constant, where the number of specific allocations belonging to each of those allocation types grows polynomially in U. Altogether this will establish the polynomial growth of $|\mathcal{Q}|$. As we are solely interested in an upper-bound $\hat{U}$ we will neglect mask constraints as these only further constrain the set of feasible allocations $\mathcal{Q}$. Feasibility of all users loading θ bits implies $r_u(\mathbf{p}) \geq \theta, \forall u \in \mathcal{U}$. Reformulating (1.3) we obtain, $\forall u \in \mathcal{U}$,

$$1 \geq (2^\theta - 1)\Gamma\Big(\sum_{i \in \mathcal{U}\backslash u} \frac{H_{ui}}{H_{uu}}\frac{p_i}{p_u} + \frac{N_u}{H_{uu}p_u}\Big). \tag{A.5}$$

Using $\frac{N_u}{H_{uu}p_u} \geq 0$, $\frac{H_{ui}}{H_{uu}} \geq \alpha > 0$, and assuming $u = \operatorname{argmin}_{i \in \mathcal{U}}\{p_i\}$, we have $p_i/p_u \geq 1$ and the necessary condition for feasibility $U \leq \hat{U}$, where $\hat{U}$ derived from (A.5) is given in (2.4). Note that for symmetric scenarios with $\frac{H_{ui}}{H_{uu}} = \alpha, N_u = 0, \forall u \in \mathcal{U}$, this bound is in fact tight as all power-allocations are equal at optimum, i.e., $p_u = p_i, \forall u, i \in \mathcal{U}$. Next we apply the method of types [43, Ch. 11.1] where any vector $\mathbf{r} = \mathbf{r}(\mathbf{p}), \mathbf{p} \in \mathcal{Q}$, is characterized by a histogram (a "type") $T_{\mathbf{r}}$ out of the set of all U-user histograms $\mathcal{T}_U$, specifying the relative number of occurrences $T_{\mathbf{r}}(k \cdot \theta)$ of any number of bits $(k \cdot \theta) \in \mathcal{B}, 0 \leq k \leq |\mathcal{B}| - 1$, in $\mathbf{r}$. As any specific number of bits can only appear U times, we have for the number of types [43, Thm. 11.1.1]

$$|\mathcal{T}_U| \leq (U + 1)^{|\mathcal{B}|} \doteq m(U). \tag{A.6}$$

The set of bit-loading sequences leading to a certain type T is denoted by its type class $\mathcal{S}(T)$ of size

$$|\mathcal{S}(T)| = \begin{pmatrix} U \\ UT(0), \ldots, UT(\hat{\theta}) \end{pmatrix} \leq 2^{UH(T)} \tag{A.7a}$$

$$\leq 2^{U(\hat{\theta}/\theta+1) \cdot \frac{[U/(\hat{\theta}/\theta+1)]}{U} \cdot \log\left(\frac{[U/(\hat{\theta}/\theta+1)]}{U}\right)} \doteq n(U), \tag{A.7b}$$

where the first inequality follows from [43, Thm. 11.1.3] and $H(\cdot)$ denotes the entropy function. Now we use the fact that interference among users limits the number of types. More precisely, we have a correspondence between a type $\hat{T} \in \mathcal{T}_{\hat{U}}$ and a type $T \in \mathcal{T}_{\hat{U}+v}$ given by

$$\mathcal{T}_{\hat{U}+v} = \left\{ \left(\frac{\hat{U}\hat{T}(0) + v}{\hat{U}+v}, \frac{\hat{U}\hat{T}(\theta)}{\hat{U}+v}, \ldots, \frac{\hat{U}\hat{T}(\hat{\theta})}{\hat{U}+v}, \right) \mid \hat{T} \in \mathcal{T}_{\hat{U}} \right\}. \tag{A.8}$$

This holds as even the type with the largest frequency of occurrence of a non-zero number of bits does not allow for further users loading a positive number of bits when $U > \hat{U}$. In other words, $|\mathcal{T}_U| = |\mathcal{T}_{\hat{U}}| \leq m(\hat{U}), \forall U \geq \hat{U}$. We will write $T^{\hat{T}}$ to denote a type in $\mathcal{T}_U$ formed from a type $\hat{T} \in \mathcal{T}_{\hat{U}}$ according to (A.8). Assuming any $U > \hat{U} + v, v > \hat{U}$, we have

$$|\mathcal{S}(T^{\hat{T}})| = \frac{U \cdot (U-1) \cdot \ldots \cdot (\hat{U}+1)}{(\hat{U}\hat{T}(0) + v) \cdot \ldots \cdot (\hat{U}\hat{T}(0) + 1)} |\mathcal{S}(\hat{T})| \tag{A.9a}$$

$$\leq \frac{U \cdot (U-1) \cdot \ldots \cdot (\hat{U}+1)}{v \cdot (v-1) \cdot \ldots \cdot 1} n(\hat{U}) \tag{A.9b}$$

$$= \frac{U \cdot (U-1) \cdot \ldots \cdot (U-\hat{U}+1)}{\hat{U}!} n(\hat{U}) \tag{A.9c}$$

$$= O\left(U^{\hat{U}}\right) \tag{A.9d}$$

where in (A.9a) we use the fact that only the number of occurrences of 0 bits grows for $U \geq \hat{U}$, in (A.9b) we use the bound in (A.7b) and bound the expression by assuming $\hat{T}(0) = 0$, and in (A.9c) we use the assumption $v > \hat{U}$. Summarizing, for any $U > \hat{U}$ we have that $|\mathcal{T}_U| \leq m(\hat{U})$ and $|\mathcal{S}(T)|$ is polynomially bounded by (A.9d), $\forall T \in \mathcal{T}_U$, concluding the proof. $\qquad \square$

A.4 Proof of Theorem 4

Proof. Our aim now is to bound the duality-gap $\zeta \doteq P^*_{(\mathbf{R},\hat{\mathbf{P}})} - D^*_{(\mathbf{R},\hat{\mathbf{P}})}$ between the dual problem in (1.10) with optimal value $D^*_{(\mathbf{R},\hat{\mathbf{P}})}$ to the primal problem in (1.8)

with optimal value $P^*_{(\mathbf{R},\hat{\mathbf{P}})}$. We start by formulating a perturbed primal problem as

$$P^*_{(\hat{\mathbf{R}},\hat{\mathbf{P}})} = \underset{\mathbf{p}^c \in \mathcal{Q}^c, c \in \mathcal{C}}{\text{minimize}} \quad \sum_{c \in \mathcal{C}} f^c(\mathbf{p}^c, \hat{\mathbf{w}}, \check{\mathbf{w}}) \tag{A.10a}$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} r^c_u(\mathbf{p}^c) \geq \hat{R}_u, \quad \sum_{c \in \mathcal{C}} p^c_u \leq \hat{P}_u, \quad \forall u \in \mathcal{U}, \tag{A.10b}$$

where $\mathcal{Q}^c$ as defined in (1.7) is the set of feasible PSD's on subcarrier c, resulting in bit-loadings for each user $u \in \mathcal{U}$ out of the discrete set $\mathcal{B}$, and where $\hat{\mathbf{R}} \in \mathbb{R}^U$ is defined in (2.17) and (2.14), respectively. Assuming a mask on power spectral densities per subcarrier and user $\hat{p}^c_u$ ensures compactness of $\mathcal{Q}^c$. We will now use the set of all possible combinations of objective values, sum-rates, and sum-powers $\check{\mathcal{Q}} = \sum_{c \in \mathcal{C}} \check{\mathcal{Q}}^c$ as defined in (2.5). Bounding the duality-gap for the primal problem in (1.8) is equivalent to bounding the minimum distance in objective value between $\check{\mathcal{Q}}$ and its convex hull conv($\check{\mathcal{Q}}$) at the given target-rates and maximum sum-powers $[\mathbf{R}^T, \hat{\mathbf{P}}^T]^T$, cf. Figure 2.4. Furthermore, the dual optimization in (1.10) can be interpreted as the search for a non-vertical hyperplane supporting $\check{\mathcal{Q}}$ and maximizing its intersection with a vertical axis at $[\mathbf{R}^T, \hat{\mathbf{P}}^T]^T$ [12]. Based thereupon, let $\bar{\mathbf{x}} = \begin{bmatrix} \bar{s}, \bar{\mathbf{r}}^T, \bar{\mathbf{p}}^T \end{bmatrix}^T \in$ conv($\check{\mathcal{Q}}$) be a point where $\bar{\mathbf{r}} = \sum_{c \in \mathcal{C}} \bar{\mathbf{r}}^c \succeq \hat{\mathbf{R}}$, $\bar{\mathbf{p}} = \sum_{c \in \mathcal{C}} \bar{\mathbf{p}}^c \preceq \hat{\mathbf{P}}$, and $\bar{s} = \sum_{c \in \mathcal{C}} \bar{s}^c = D^*(\hat{\mathbf{R}}, \hat{\mathbf{P}})$, and where $[\bar{s}^c, (\bar{\mathbf{r}}^c)^T, (\bar{\mathbf{p}}^c)^T]^T \in$ conv($\check{\mathcal{Q}}^c$), $D^*_{(\hat{\mathbf{R}},\hat{\mathbf{P}})}$ being the optimal cost of the partial dual problem to (A.10) after relaxing the coupling constraints in (A.10b). As in [11, Sec. 5.6.1], we apply the Shapley-Folkman theorem, by which for any point $\mathbf{x} \in$ conv($\check{\mathcal{Q}}$) it holds that

$$\mathbf{x} \in \left[\sum_{c \in \mathcal{C} \setminus \mathcal{D}(\mathbf{x})} \check{\mathcal{Q}}^c + \sum_{c \in \mathcal{D}(\mathbf{x})} \text{conv}\left(\check{\mathcal{Q}}^c\right) \right], \tag{A.11}$$

where $\mathcal{D}(\mathbf{x}) \subseteq \mathcal{C}$, $\mid \mathcal{D}(\mathbf{x}) \mid = (U+1)$. Hence, applying this theorem to the previously defined point $\bar{\mathbf{x}}$, we may write

$$\bar{\mathbf{r}} = \sum_{c \in \mathcal{C} \setminus \mathcal{D}(\bar{\mathbf{x}})} \mathbf{r}^c(\bar{\mathbf{p}}^c) + \sum_{c \in \mathcal{D}(\bar{\mathbf{x}})} \bar{\mathbf{r}}^c \qquad \succeq \hat{\mathbf{R}}, \tag{A.12a}$$

$$\bar{\mathbf{p}} = \sum_{c \in \mathcal{C} \setminus \mathcal{D}(\bar{\mathbf{x}})} \bar{\mathbf{p}}^c + \sum_{c \in \mathcal{D}(\bar{\mathbf{x}})} \bar{\mathbf{p}}^c \qquad \preceq \hat{\mathbf{P}}, \tag{A.12b}$$

$$\bar{s} = \sum_{c \in \mathcal{C} \setminus \mathcal{D}(\bar{\mathbf{x}})} f^c(\bar{\mathbf{p}}^c, \hat{\mathbf{w}}, \check{\mathbf{w}}) + \sum_{c \in \mathcal{D}(\bar{\mathbf{x}})} \bar{s}^c = D^*(\hat{\mathbf{R}}), \tag{A.12c}$$

where $\mathbf{r}^c(\bar{\mathbf{p}}^c) = \bar{\mathbf{r}}^c$ and $\bar{\mathbf{p}}^c \in \mathcal{Q}^c$, $\forall c \in \mathcal{C} \setminus \mathcal{D}(\bar{\mathbf{x}})$. Note that only the second part in all three equations in (A.12) involves points out of conv($\check{\mathcal{Q}}^c$) for subcarriers $c \in \mathcal{D}(\bar{\mathbf{x}})$. Bounding $\bar{r}^c_u \leq \max_{\mathbf{p}^c \in \mathcal{Q}^c}\{r^c_u(\mathbf{p}^c)\}$ on subcarriers $c \in \mathcal{D}(\bar{\mathbf{x}})$, noting that $\bar{\mathbf{r}}^c \succeq \mathbf{0}$, and

using the initial definition of $\hat{\mathbf{R}}$ in (2.17) and (2.14), from Equation (A.12a) we infer

$$\sum_{c\in\mathcal{C}} \mathbf{r}^c\left(\bar{\mathbf{p}}^c\right) \succeq \mathbf{R}. \tag{A.13}$$

Note that, without making any assumptions on the set $\mathcal{D}(\bar{\mathbf{x}})$, it was necessary to take the highest bit-loadings in (2.14) as the rate constraints have to hold for each user. Inequality (A.13) holds for any $\bar{\mathbf{p}}^c \in \mathcal{Q}^c, \forall c \in \mathcal{D}(\bar{\mathbf{x}})$, and we may choose to set the powers on these $(U+1)$ subcarriers to zero. By (A.13), feasibility in (A.12b) by the choice of $\bar{\mathbf{p}}^c, \forall c \in \mathcal{D}(\bar{\mathbf{x}})$, and feasibility of $\bar{\mathbf{p}}$ with respect to the per-subcarrier constraints in (1.7), the thereby created $\bar{\mathbf{p}}$ is shown to be feasible for the original, unperturbed problem in (1.8) with target-rates $\mathbf{R}$, maximum sum-power $\hat{\mathbf{P}}$, and optimal cost $P^*_{(\mathbf{R},\hat{\mathbf{P}})}$. Hence, Theorem 4 follows from (cf. Figure 2.4)

$$D^*(\mathbf{R}) \leq P^*(\mathbf{R}) \leq \sum_{c\in\mathcal{C}} f^c(\bar{\mathbf{p}}^c, \hat{\mathbf{w}}, \breve{\mathbf{w}}) \leq D^*(\hat{\mathbf{R}}) + \sum_{u\in\mathcal{U}} \breve{w}_u \cdot \Delta R_u, \tag{A.14}$$

where the first inequality holds due to weak-duality relation in (1.13), the second inequality holds since $\bar{\mathbf{p}}$ is not necessarily the optimum of the primal problem (1.8), and the third inequality holds due to the above choice $\bar{\mathbf{p}}^c = \mathbf{0}$ and $\bar{s}^c \geq -\sum_{u\in\mathcal{U}} \breve{w}_u \Delta R_u$, $\forall c \in \mathcal{D}(\bar{\mathbf{x}})$, cf. (A.12c). $\qquad\square$

A.5 Proof of Theorem 5

Proof. We aim now to proof Theorem 5 by modification of the bound in (2.16) for the special case of a sum-power minimization problem, i.e., $\breve{\mathbf{w}} = \mathbf{0}$. As a first step we modify (2.16) into a bound formulated in terms of primal objective values. This is necessary as we want to work with feasible integer bit-loadings and avoid the obligation to consider convex combinations of those, cf. the proof of Theorem 4. Define $D^*_{(\hat{\mathbf{R}},\hat{\mathbf{P}})}$ $(P^*_{(\hat{\mathbf{R}},\hat{\mathbf{P}})})$ and $D^*_{(\breve{\mathbf{R}},\hat{\mathbf{P}})}$ $(P^*_{(\breve{\mathbf{R}},\hat{\mathbf{P}})})$ as the optimal objective values to the dual problem in (1.10) (primal problem in (1.8)), with modified target-rates $\hat{R}_u$ as in (2.17) and

$$\breve{R}_u = \max\{0, R_u - \Delta R_u\}, \ \forall u \in \mathcal{U}, \tag{A.15}$$

where ΔR_u is defined in (2.14), respectively. Note that all these modified problems are feasible by the assumptions of the theorem. Then we have

$$P^*_{(\hat{\mathbf{R}},\hat{\mathbf{P}})} \geq D^*_{(\hat{\mathbf{R}},\hat{\mathbf{P}})} \geq P^*_{(\mathbf{R},\hat{\mathbf{P}})} \geq D^*_{(\mathbf{R},\hat{\mathbf{P}})}, \tag{A.16}$$

where the second inequality follows from (2.16) and the other inequalities from the weak-duality relation in (1.13). If $R_u \geq \Delta R_u$ we may similarly apply (2.16) to show

$$D^*_{(\mathbf{R},\hat{\mathbf{P}})} \geq P^*_{(\check{\mathbf{R}},\hat{\mathbf{P}})} - \sum_{u \in \mathcal{U}} \check{w}_u \Delta R_u. \tag{A.17}$$

It follows from the proof of Theorem 4 that (A.17) also holds in case of $R_u < \Delta R_u$ and $\check{\mathbf{R}}$ chosen as in (A.15). More specifically, we see from (A.12a) and the non-negativity of rates $r^c_u(\cdot)$ that (A.14) remains valid also in this case. Combining (A.16), (A.17), and (2.16) we have

$$\zeta \leq P^*(\hat{\mathbf{R}}) - P^*(\check{\mathbf{R}}), \tag{A.18}$$

where $\hat{\boldsymbol{R}} - \check{\boldsymbol{R}} \preceq 2 \cdot \boldsymbol{\Delta R}$. Therefore we proceed to bound ζ by bounding the extra power necessary for each user u to load $2 \cdot \Delta R_u$ bits. Consider the worst-case cost Δp^c_u of loading a bit-step for user u on subcarrier c as defined in (2.20). For example, it also holds that

$$\varepsilon \cdot \Delta p^c_u \geq \max_{\substack{\{\{\mathbf{p}^c, \tilde{\mathbf{p}}^c\} \mid \mathbf{p}^c, \tilde{\mathbf{p}}^c \in \mathcal{Q}^c, \\ r^c_u(\tilde{\mathbf{p}}^c) = r^c_u(\mathbf{p}^c) + \varepsilon \cdot \theta, \\ r^c_i(\tilde{\mathbf{p}}^c) = r^c_i(\mathbf{p}^c), \forall i \in \mathcal{U} \setminus u\}}} \left\{ \sum_{u \in \mathcal{U}} \hat{w}_u \left(\tilde{p}^c_u - p^c_u \right) \right\}, \tag{A.19}$$

as loading ε bit-steps can be regarded as sequentially loading ε times one bit-step. Hence, no matter how many of the extra bit-steps are loaded jointly, the extra power-cost the $2 \cdot \Delta R_u$ bits incur is bounded by $2 \cdot \Delta R_u/\theta$ times the maximum cost over subcarriers Δp^c_u of loading one bit-step. Concluding, the bound (2.19) follows from (A.18) and

$$P^*(\hat{\mathbf{R}}) - P^*(\check{\mathbf{R}}) \leq 2 \cdot \sum_{u \in \mathcal{U}} \Delta R_u/\theta \cdot \max_{c \in \mathcal{C}} \{\Delta p^c_u\}. \tag{A.20}$$

We notably avoided making any assumptions on the optimal power-allocation in the subproblems in (1.12) under target-rates $\check{\mathbf{R}}$ yielding the optimal objective $P^*_{(\check{\mathbf{R}},\hat{\mathbf{P}})}$ by taking the *maximum* cost of loading a bit-step. An idea to strengthen the bound is for instance to assume that from the optimal power-allocation of (1.8) with decreased target-rates $\check{\mathbf{R}}$ every user u can load an extra bit on $2\Delta R_u$ subcarriers. The resulting bound would be lower, especially when subcarriers differ significantly and the number of users increases. $\qquad \square$

A.6 Proof of Theorem 7

Proof. For $U = 1$ and arbitrary C the objective in (2.24a) is simply a single non-linear, monotonously increasing function (a square-root) of the user's sum-power, and omitting this function does therefore not change the optimum of the problem in (2.24) [22], yielding an identical formulation as of the transmit-power minimization-problem in (2.23). In the case of $C = 1$ and arbitrary U the target-rates in (1.8b) uniquely define the minimal per-user transmit powers necessary to support the target-rates [194]. However, as the LDP model in (2.21) as a function of the per-user transmit sum-power is monotonously increasing, any other power-allocation feasible in (2.24b) than this minimal one would have a higher LDP consumption, and the minimum TP solution for the problem in (2.23) is therefore also optimal in the LDP minimization problem in (2.24). □

A.7 Solution of the Relaxation in (3.4)

In the following we detail the exact solution of the problem in (3.4) and use the short notation $\lfloor r_u^c(\mathbf{p}^c) \rfloor_\theta$ and $\lceil r_u^c(\mathbf{p}^c) \rceil_\theta$ to denote the rate rounded up or down to the next integer multiple of the bit-step θ, respectively. The problem in (3.4) is separable among users $i, (u+1) \le i \le U$. The continuous relaxation of each of those separated problems is convex and given as

$$\underset{r_i \in [0, \hat{\theta}_i^{\mathrm{mod}}]}{\mathrm{minimize}} \; \tilde{f}_i(r_i) = (\hat{w}_i + \nu_i)(2^{r_i} - 1)\frac{\Gamma \tilde{N}_i}{H_{ii}} - (\breve{w}_i + \lambda_i)r_i, \tag{A.21}$$

where $\hat{\theta}_i^{\mathrm{mod}} = \min\{\hat{\theta}, \lfloor \log_2(1 + \frac{H_{ii}\hat{p}_i}{\Gamma \tilde{N}_i}) \rfloor_\theta\}$ and we use the fact that the constraints in (3.4b) hold with equality at optimum to replace the variable p_i. The problem in (A.21) can be given analytically, for each remaining user $i, (u + 1) \le i \le U$ independently, using first-order optimality conditions $\partial \tilde{f}_i(r_i)/\partial r_i = 0$, leading to

$$r_i^{\mathrm{tmp}} = \min\left\{ \log_2\left(\frac{\lambda_u H_{ii}}{(\hat{w}_i + \nu_i)\log(2)\tilde{N}_i \Gamma} \right), \hat{\theta}_i^{\mathrm{mod}} \right\}, \quad (u+1) \le i \le U, \tag{A.22}$$

with corresponding power-allocation

$$p_i^{\mathrm{tmp}}(r_i^{\mathrm{tmp}}) = (2^{r_i^{\mathrm{tmp}}} - 1)\frac{\Gamma \tilde{N}_i}{H_{ii}}, \quad (u+1) \le i \le U. \tag{A.23}$$

Inserting (A.22) in (A.23) we can interpret $p_i^{\mathrm{tmp}}(r_i^{\mathrm{tmp}})$ as the water-filling solution [22, Ex. 5.2] for the water-level $(\hat{w}_i + \nu_i)/((\breve{w}_i + \lambda_i)\log(2))$ under bit-cap and PSD mask constraints, respectively. The solution $\bar{r}_i, \bar{p}_i, (u+1) \le i \le U$, of the discrete problem in (3.4) can then be computed by rounding the user's rates r_i^{tmp} to one of the two

nearest integer multiples of θ with the lower objective, i.e.,

$$[\bar{r}_i, \bar{p}_i] = \underset{\substack{\{r_i \in \{\lfloor r_i^{\mathrm{tmp}} \rfloor_\theta, \lceil r_i^{\mathrm{tmp}} \rceil_\theta\}, \\ p_i = p_i^{\mathrm{tmp}}(r_i) | r_i \le \hat{\theta}_i^{\mathrm{mod}}\}}}{\operatorname{argmin}} \left\{ (\hat{w}_i + \nu_i) p_i - (\breve{w}_i + \lambda_i) r_i \right\}. \qquad (A.24)$$

This holds due to the aforementioned convexity and user-independence.

A.8 Proof of Theorem 8

Proof. Assume feasible weights $\tilde{\xi}', \tilde{\xi}_i^c, i \in \mathcal{I}^{c,(k)}, c \in \mathcal{C}$, in (4.1) with $0 < \tilde{\xi}' < 1$. Consider weights $\xi' = 0, \xi_i^c = \tilde{\xi}_i^c/(1 - \tilde{\xi}'), \forall i \in \mathcal{I}^{c,(k)}$, for the problem in (4.1), setting the weights of columns not included in (4.1) to zero, i.e., $\xi_i^c = 0, \forall i \in \mathcal{I}^c \setminus \mathcal{I}^{c,(k)}, \forall c \in \mathcal{C}$. Regarding (4.1b)–(4.1d), from the feasibility of $\tilde{\xi}', \{\tilde{\xi}_i^c\}$ it follows that the weights $\xi', \{\xi_i^c\}$ are also feasible in (2.3) and (4.1). Using (1.9) we can write the objective value for weights $\xi', \{\xi_i^c\}$ in (4.1a) as

$$\sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}^{c,(k)}} \left(\hat{\mathbf{w}}^T \mathbf{p}^{c,i} - \breve{\mathbf{w}}^T \mathbf{r}^c(\mathbf{p}^{c,i}) \right) \xi_i^c \qquad (A.25a)$$

$$\preceq \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}^{c,(k)}} \left(\hat{\mathbf{w}}^T \mathbf{p}^{c,i} - \breve{\mathbf{w}}^T \mathbf{r}^c(\mathbf{p}^{c,i}) \right) \tilde{\xi}_i^c + \left(\hat{\mathbf{w}}^T \hat{\mathbf{P}} - \breve{\mathbf{w}}^T \mathbf{R} \right) \tilde{\xi}', \quad (A.25b)$$

$$\prec \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}^{c,(k)}} \left(\hat{\mathbf{w}}^T \mathbf{p}^{c,i} - \breve{\mathbf{w}}^T \mathbf{r}^c(\mathbf{p}^{c,i}) \right) \tilde{\xi}_i^c + f'(\hat{\mathbf{P}}, \mathbf{R}) \tilde{\xi}', \qquad (A.25c)$$

where the two parts in (A.25b) are obtained by multiplying (A.25a) once by $(1 - \tilde{\xi}')$ and once by $\tilde{\xi}'$, respectively, and using the definition of ξ_i^c above as well as feasibility in (4.1). Inequality (A.25c) follows as $\boldsymbol{\delta} \succ \mathbf{0}$. This implies weights $0 < \tilde{\xi}' < 1$ are not optimal in (4.1). Again from $\boldsymbol{\delta} \succ \mathbf{0}$ it follows that if weights as described in the theorem exist they will be feasible in (4.1) with $\tilde{\xi}' = 0$ and have a lower objective than $f'(\hat{\mathbf{P}}, \mathbf{R})$, concluding $\tilde{\xi}' = 0$ at optimum of (4.1). On the other hand, if no weights as described in the theorem exist we must have $\xi' > 0$ at optimum of (4.1) and $0 < \xi' < 1$ would lead to a contradiction by the arguments above, and hence $\xi' = 1$. $\qquad \square$

A.9 Alternative Derivation of the Semidefinite Relaxation based Coverage Maximization Problem in (5.5)

We can rewrite the relation $b_u^2 = b_u, u \in \mathcal{U}$, as $\operatorname{diag}(\boldsymbol{b}\boldsymbol{b}^T) = \boldsymbol{b}$ [102]. Linearizing the quadratic terms by introducing new variables $B_{uv} = b_u b_v, u, v \in \mathcal{U}$, we can write the

equivalent "lifted" problem to that in (5.1) as

$$\underset{\substack{\mathbf{p}^c \in \mathcal{Q}^c, c \in \mathcal{C}, \\ \boldsymbol{b} \in \mathbb{R}^U, \boldsymbol{B} \in \mathbb{R}^{U \times U}}}{\text{maximize}} \quad \sum_{u \in \mathcal{U}} b_u \tag{A.26a}$$

$$\text{subject to} \quad \text{diag}(\boldsymbol{B}) = \boldsymbol{b} \tag{A.26b}$$

$$\boldsymbol{B} - \boldsymbol{b}\boldsymbol{b}^T = \boldsymbol{0}, \tag{A.26c}$$

$$\text{Constraints in (5.1b) and (5.1c).} \tag{A.26d}$$

The constraint in (A.26c) is nonconvex and at this point relaxed to the convex constraint $\boldsymbol{B} - \boldsymbol{b}\boldsymbol{b}^T \succeq \boldsymbol{0}$ [102]. We can regard this relaxed constraint as the Schur complement [22, p.650] of the equivalent constraint

$$\breve{\boldsymbol{B}} = \begin{bmatrix} 1 & \boldsymbol{b}^T \\ \boldsymbol{b} & \boldsymbol{B} \end{bmatrix} \succeq \boldsymbol{0}. \tag{A.27}$$

Reversing the row and column indices does not change this condition as seen by the definition of semidefiniteness [2], leading to the formulation of a semidefinite programming (SDP) relaxation of the problem in (5.1) in the so called "diagonal form" [102], given by

$$P_{(\mathbf{R},\hat{\boldsymbol{P}})}^{cov,sdp} = \underset{\substack{\mathbf{p}^c \in \mathcal{Q}^c, c \in \mathcal{C}, \\ \breve{\boldsymbol{B}} \in \mathbb{R}^{(U+1) \times (U+1)}}}{\text{maximize}} \quad \sum_{u \in \mathcal{U}} \breve{B}_{uu} \tag{A.28a}$$

$$\text{subject to} \quad \breve{B}_{uu} = \breve{B}_{(U+1)u}, \forall u \in \mathcal{U}, \ \breve{B}_{(U+1)(U+1)} = 1, \tag{A.28b}$$

$$\breve{\boldsymbol{B}} \succeq \boldsymbol{0}, \tag{A.28c}$$

$$\sum_{c \in \mathcal{C}} r_u^c(\mathbf{p}^c) \geq R_u + \breve{B}_{uu}(\hat{R} - R_u), \quad \forall u \in \mathcal{U}, \tag{A.28d}$$

$$\text{Constraint in (5.1c),} \tag{A.28e}$$

where the additional constraint $\breve{B}_{(U+1)(U+1)} = 1$ is indeed needed to ensure that $\breve{B}_{uu} \in [0,1], u \in \mathcal{U}$, as shown in the proof of the following corollary which compares the strength of this relaxation to the linear relaxation in (5.2).

Corollary 6. *We have that $P_{(\mathbf{R},\hat{\boldsymbol{P}})}^{cov,sdp} \leq P_{(\mathbf{R},\hat{\boldsymbol{P}})}^{cov,lr}$, i.e., the SDP relaxation is never less tight than the linear relaxation. Furthermore, in the special case where the solution of the linear relaxation in (5.2) satisfies $\boldsymbol{b} \succ \boldsymbol{0}, \sum_{u \in \mathcal{U}} b_u \leq 1$, the two relaxations are in fact equivalent in the sense that $P_{(\mathbf{R},\hat{\boldsymbol{P}})}^{cov,sdp} = P_{(\mathbf{R},\hat{\boldsymbol{P}})}^{cov,lr}$ holds.*

Proof. Theorem 11 proves the corresponding statements for the case with time-sharing relaxation and the corollary follows along the arguments given in the proof of Theorem 11 by just exchanging the role of time-sharing weights $\xi_i^c, i \in \mathcal{I}^{c,(k)}, c \in \mathcal{C}$, with that of the power-allocation variables $\mathbf{p}^c, c \in \mathcal{C}$. $\qquad\square$

To make the power-allocation part in (A.28) tractable we again apply a time-sharing relaxation to the per-subcarrier power and rate allocation, leading to a modification of the sum-terms of per-subcarrier rates and power in (A.28d) and (A.28e), respectively, and the time-sharing relaxation of the SDP relaxation in the form of the problem in (5.5).

A.10 Proof of Theorem 10

Proof. To prove the first statement we show the equivalence between the Lagrange-duals of the problems in (5.4) and (5.5), respectively. We begin with the dual problem of the original binary problem in (5.4), given as

$$
\begin{aligned}
\underset{\boldsymbol{\lambda}\in\mathbb{R}_+^U,\boldsymbol{\nu}\in\mathbb{R}_+^U,\boldsymbol{\kappa}\in\mathbb{R}^U}{\text{minimize}} \; \underset{\mathbf{p}^c\in\mathcal{Q}^c,c\in\mathcal{C}}{\max} \Bigg\{ \sum_{u\in\mathcal{U}}\Big(\lambda_u\sum_{c\in\mathcal{C}}r_u^c\left(\mathbf{p}^c\right) - \nu_u\sum_{c\in\mathcal{C}}p_u^c\Big)\Bigg\} - \sum_{u\in\mathcal{U}}\lambda_u R_u \\
+ \sum_{u\in\mathcal{U}}\nu_u\hat{P}_u + \underset{\boldsymbol{b}}{\max}\Bigg\{\sum_{u\in\mathcal{U}}\big(b_u^2 + \kappa_u(b_u^2 - b_u) - \lambda_u b_u^2 R_u^+\big)\Bigg\}, \quad \text{(A.29)}
\end{aligned}
$$

where $R_u^+ = (\hat{R} - R_u)$. Clearly, the second maximization in (A.29) can be solved analytically for each user separately, where we denote the optimal objective of user u in this maximization by $f(\kappa_u, \lambda_u R_u^+)$, i.e., as a function of the dual variables κ_u and λ_u, $u \in \mathcal{U}$. By inspection we see that for this quadratic maximization problem to have a bounded objective, the following two criteria need to be met

$$
-\big(1 + \kappa_u - \lambda_u R_u^+\big) \geq 0, \quad \forall u \in \mathcal{U}, \tag{A.30a}
$$

$$
\big(1 - i(\kappa_u, \lambda_u R_u^+)\big)\kappa_u = 0, \quad \forall u \in \mathcal{U}, \tag{A.30b}
$$

where $i(x, y)$ is an indicator function giving 0 in case $(1 + x - y) = 0$, and 1 otherwise. While (A.30a) ensures concavity, the complementarity condition in (A.30b) captures the case when $i(\kappa_u, \lambda_u R_u^+)$ is zero and the maximization is hence a linear problem with objective $-\kappa_u b_u$, making it necessary to enforce $\kappa_u = 0$ to ensure boundedness of the objective. In case these criteria hold, the second (quadratic and unconstrained) maximization problem in (A.29) is solvable by first-order optimality conditions, and its optimal objective value given by

$$
f(\kappa_u, \lambda_u R_u^+) = \begin{cases} -\frac{1}{4}\kappa_u^2 \frac{1}{1+\kappa_u-\lambda_u R_u^+} & \text{if } i(\kappa_u, \lambda_u R_u^+) == 1, \\ 0 & \text{otherwise,} \end{cases} \quad \forall u \in \mathcal{U}. \tag{A.31}
$$

Next we introduce an auxiliary variable $-\tilde{\kappa}$ for the just computed optimal objective and rewrite the problem in (A.29) as

$$
\begin{aligned}
\underset{\boldsymbol{\lambda}\in\mathbb{R}_+^U, \boldsymbol{\nu}\in\mathbb{R}_+^U, \boldsymbol{\kappa}\in\mathbb{R}^U, \tilde{\kappa}\in\mathbb{R}}{\text{minimize}} \quad & \underset{\mathbf{p}^c\in\mathcal{Q}^c, c\in\mathcal{C}}{\max}\Big\{ \sum_{u\in\mathcal{U}}(\lambda_u \sum_{c\in\mathcal{C}} r_u^c\,(\mathbf{p}^c) - \nu_u \sum_{c\in\mathcal{C}} p_u^c)\Big\} - \sum_{u\in\mathcal{U}} \lambda_u R_u \\
& + \sum_{u\in\mathcal{U}} \nu_u \hat{P}_u - \tilde{\kappa} & \text{(A.32a)}
\end{aligned}
$$

$$
\text{subject to} \quad -\tilde{\kappa} - \sum_{u\in\mathcal{U}} f(\kappa_u, \lambda_u R_u^+) \geq 0, \tag{A.32b}
$$

$$
\text{Constraints (A.30a) and (A.30b).} \tag{A.32c}
$$

We recognize the constraints in (A.32b) and (A.32c) as the generalized Schur complement conditions for positive semidefiniteness of singular block-matrices [22, p. 651], leading to the equivalent formulation of the problems in (A.29) and (A.32) as

$$
\begin{aligned}
\underset{\boldsymbol{\lambda}\in\mathbb{R}_+^U, \boldsymbol{\nu}\in\mathbb{R}_+^U, \boldsymbol{\kappa}\in\mathbb{R}^U, \tilde{\kappa}\in\mathbb{R}}{\text{minimize}} \quad & \underset{\mathbf{p}^c\in\mathcal{Q}^c, c\in\mathcal{C}}{\max}\Big\{ \sum_{u\in\mathcal{U}}(\lambda_u \sum_{c\in\mathcal{C}} r_u^c\,(\mathbf{p}^c) - \nu_u \sum_{c\in\mathcal{C}} p_u^c)\Big\} - \sum_{u\in\mathcal{U}} \lambda_u R_u \\
& + \sum_{u\in\mathcal{U}} \nu_u \hat{P}_u - \tilde{\kappa} & \text{(A.33a)}
\end{aligned}
$$

$$
\text{subject to} \quad \begin{bmatrix} \operatorname{diag}\left(-\mathbf{1} - \boldsymbol{\kappa} + \operatorname{diag}(\mathbf{R}^+)\boldsymbol{\lambda}\right) & \frac{1}{2}\boldsymbol{\kappa} \\ \frac{1}{2}\boldsymbol{\kappa}^T & -\tilde{\kappa} \end{bmatrix} \succeq \mathbf{0}. \tag{A.33b}
$$

We continue by comparing this problem to the Lagrange-dual of the SDP [102] in (5.5). Due to symmetry of $\breve{\boldsymbol{B}}$ the constraint in (5.5b) can be equivalently written as $\breve{B}_{uu} = 1/2(\breve{B}_{(U+1)u} + \breve{B}_{u(U+1)}), \forall u \in \mathcal{U}$. Using this modification and $\boldsymbol{\epsilon}, \boldsymbol{\kappa}, \tilde{\kappa}, \mathbf{Z}, \boldsymbol{\lambda}, \boldsymbol{\nu}$, and $\boldsymbol{\Phi}$ to denote the Lagrange multipliers of the constraints in (5.5a)–(5.5e), respectively, we write the dual SDP to the problem in (5.5) as

$$
\begin{aligned}
\underset{\substack{\boldsymbol{\lambda}\in\mathbb{R}_+^U, \boldsymbol{\nu}\in\mathbb{R}_+^U, \boldsymbol{\kappa}\in\mathbb{R}^U, \tilde{\kappa}\in\mathbb{R}, \\ \boldsymbol{\Phi}\in\mathbb{R}^C, \boldsymbol{\epsilon}\in\mathbb{R}^{\sum_{c\in\mathcal{C}} |i\in\mathcal{I}^c|}, \\ \mathbf{Z}\in\mathbb{R}^{(U+1)\times(U+1)}, \mathbf{Z}\succeq\mathbf{0}}}{\text{minimize}} \quad & \underset{\substack{\xi_i^c, i\in\mathcal{I}^c, c\in\mathcal{C}, \\ \breve{\boldsymbol{B}}\in\mathbb{R}^{(U+1)\times(U+1)}}}{\max} \quad \sum_{i\in\mathcal{I}^c, c\in\mathcal{C}} \xi_i^c\Big(\sum_{u\in\mathcal{U}}(\lambda_u r_u^c(\mathbf{p}^{c,i}) - \nu_u p_u^{c,i})
\end{aligned}
$$

$$
\begin{aligned}
&-\Phi_c + \epsilon^{c,i}\Big) + \sum_{u\in\mathcal{U}}\Big(-\lambda_u R_u + \nu_u \hat{P}_u\Big) + \sum_{c\in\mathcal{C}} \Phi_c - \tilde{\kappa} + \tilde{\kappa}\breve{B}_{(U+1)(U+1)} \\
&+ \sum_{u,\mathrm{v}\in\mathcal{U}} Z_{u\mathrm{v}}\breve{B}_{u\mathrm{v}} + \sum_{u\in\mathcal{U}}\Big(\breve{B}_{uu} - \lambda_u R_u^+ \breve{B}_{uu} + \kappa_u(\breve{B}_{uu} - \frac{1}{2}\breve{B}_{(U+1)u} - \frac{1}{2}\breve{B}_{u(U+1)})\Big)
\end{aligned} \tag{A.34}
$$

The (unconstrained) maximization in (A.34) targets two linear terms, namely a term of the form $\sum_{i\in\mathcal{I}^c, c\in\mathcal{C}} A^{c,i}\xi_i^c$ and a separate term of the form $\operatorname{tr}(\mathbf{M}\breve{\boldsymbol{B}})$, where $\operatorname{tr}(\cdot)$ denotes the trace of a matrix, $A^{c,i} \in \mathbb{R}, i \in \mathcal{I}^c, c \in \mathcal{C}$, and $\mathbf{M} \in \mathbb{R}^{(U+1)\times(U+1)}$. Hence, for this maximization to be bounded we can introduce the constraints $\mathbf{M} = \mathbf{0}$ and

$A^{c,i} = 0, i \in \mathcal{I}^c, c \in \mathcal{C}$. Dropping the positive and positive semidefinite variables $\epsilon^{c,i}, i \in \mathcal{I}^c, c \in \mathcal{C}$, and $\boldsymbol{Z}$, respectively, from these equations we obtain inequalities, leading to an equivalent formulation of (A.34) given as

$$\underset{\substack{\boldsymbol{\lambda} \in \mathbb{R}_+^U, \boldsymbol{\nu} \in \mathbb{R}_+^U, \\ \boldsymbol{\kappa} \in \mathbb{R}^U, \tilde{\kappa} \in \mathbb{R}, \boldsymbol{\Phi} \in \mathbb{R}^C}}{\text{minimize}} \quad \sum_{u \in \mathcal{U}} \left(-\lambda_u R_u + \nu_u \hat{P}_u \right) + \sum_{c \in \mathcal{C}} \Phi_c - \tilde{\kappa} \tag{A.35a}$$

$$\text{subject to} \quad \sum_{u \in \mathcal{U}} (\lambda_u r_u^c(\mathbf{p}^{c,i}) - \nu_u p_u^{c,i}) \leq \Phi_c, \quad i \in \mathcal{I}^c, c \in \mathcal{C}, \tag{A.35b}$$

$$\begin{bmatrix} \text{diag} \left(-\mathbf{1} - \boldsymbol{\kappa} + \text{diag}(\mathbf{R}^+)\boldsymbol{\lambda} \right) & \frac{1}{2}\boldsymbol{\kappa} \\ \frac{1}{2}\boldsymbol{\kappa}^T & -\tilde{\kappa} \end{bmatrix} \succeq \mathbf{0}. \tag{A.35c}$$

Now we can use the same argument as in the proof of Theorem 1 in Appendix A.1: As we minimize over $\Phi_c, c \in \mathcal{C}$, in (A.35a) and these variables are lower-bounded in (A.35b), we can replace them by the maximum lower-bound, leading to an objective as in (A.33a). At this point we have that the problems in (A.35) and (A.33) are equivalent and therefore that the problems in (5.5) and (5.4) are Lagrange bi-duals of each other.

The second statement of the theorem follows from Slater's condition for constraint qualification [22, p.226] which ensures strong duality for convex problems, such as the SDP in (5.5). More precisely, the conditions are the feasibility of all affine inequalities and the strong feasibility (i.e., with strict inequality) of all other convex inequalities. The only non-affine constraint of the SDP in (5.5) is the positive semidefiniteness constraint in (5.5c). If a solution with $\breve{B}_{uu} > 0, u \in \mathcal{U}$, satisfies the constraints in (5.5d) as given by the assumptions of the theorem, then a solution $\breve{\boldsymbol{B}} \succ \mathbf{0}$ strictly satisfying the constraint in (5.5c) is easily found as shown in the proof of Theorem 11 in Appendix A.11, concluding the proof. □

A.11 Proof of Theorem 11

Proof. To prove the first statement it suffices to show that any feasible solution in (5.5) can be directly mapped to a feasible solution in (5.3) and equivalence of objective values. The positive semidefinite constraint in (5.5c) implies by definition that $\mathbf{x}^T \breve{\boldsymbol{B}} \mathbf{x} \geq 0, \forall \mathbf{x} \neq \mathbf{0}$. By specific choices of $\mathbf{x}$ it follows that $\breve{B}_{uu} \geq 0$ and $\breve{B}_{uv}^2 \leq \breve{B}_{uu} \breve{B}_{vv}, \forall u \in \mathcal{U}, v \in \mathcal{U} \setminus \{u\}$, cf. [2, Ex. 8.7]. Hence, together with the constraint in (5.5b) it holds that $0 \leq \breve{B}_{uu} \leq 1, \forall u \in \mathcal{U}$. Therefore, considering the specific choice of variables $b_u = \breve{B}_{uu}, u \in \mathcal{U}$, together with identical time-sharing weights $\xi_i^c, i \in \mathcal{I}^c, c \in \mathcal{C}$, we see that all the constraints in (5.3) are satisfied with identical objective of the corresponding variables $\breve{\boldsymbol{B}}$ in the SDP relaxation in (5.5), proving the first part of the theorem. Conversely, for the second part we additionally show that the specific conditions on the optimum $\boldsymbol{b}$ of the linear relaxation in (5.3)

imply feasibility of the specific choice $\breve{\boldsymbol{B}} = \begin{bmatrix} \operatorname{diag}(\boldsymbol{b}) & \boldsymbol{b} \\ \boldsymbol{b}^T & 1 \end{bmatrix}$, for the constraints in (5.5).

More precisely, the constraint in (5.5b) is satisfied by the choice of $\breve{\boldsymbol{B}}$, leaving only to show that constraint (5.5c) holds, which by appropriate change of the row and column index order and the Schur complement conditions [22, p. 650] is valid if and only if $\left(\operatorname{diag}(\boldsymbol{b}) - \boldsymbol{b}\boldsymbol{b}^T\right) \succeq \mathbf{0}$. By [2, Ex. 8.43] this relation holds, given $\boldsymbol{b} \succ \mathbf{0}$ as ensured by the assumptions of the theorem, if and only if $\boldsymbol{b}^T \operatorname{diag}(\boldsymbol{b})^{-1}\boldsymbol{b} \leq 1$, which is again satisfied by the given assumption $\sum_{u \in \mathcal{U}} b_u \leq 1$. Feasibility of the sum-rate and sum-power constraints in (5.5d) and (5.5e), respectively, follow by taking identical time-sharing weights as at the optimum of (5.3). The objective value of the chosen $\breve{\boldsymbol{B}}$ in (5.5a) is identical to that of the optimal variables $\boldsymbol{b}$ in (5.3a), which concludes the proof. $\qquad\square$

A.12 Proof of Theorem 13

In the following we derive feasibility conditions for robust solutions of the problem in (6.13). These conditions dependent on the matrix $\mathbf{F}^c, c \in \mathcal{C}$, defined in Section 1.2.3, and the multiplicative uncertainty parameters $\boldsymbol{\varepsilon} \in \mathbb{R}^U$. Regarding the feasibility condition in Equation (1.6), for feasibility of (6.13) it has to hold that the spectral radius $\rho\left(\mathbf{F}^c + \Delta\mathbf{F}^{c,box}\right) < 1$, where $\Delta\mathbf{F}^{c,box} \in \mathbb{R}^{U \times U}, \forall c \in \mathcal{C}$,

$$\Delta F_{u\text{v}}^{c,box} = \Gamma\gamma_u^c\Lambda_{u\text{v}}^c H_{u\text{v}}^c / H_{uu}^c = \Lambda_{u\text{v}}^c\mathbf{F}_{u\text{v}}^c \leq \varepsilon_u\mathbf{F}_{u\text{v}}^c, \tag{A.36}$$

and $\Lambda_{u\text{v}}^c$ are arbitrary values with $\Lambda_{u\text{v}}^c \leq \varepsilon_u$, $\forall u \in \mathcal{U}, c \in \mathcal{C}$, cf. the constraints in (6.13) and the definitions in (6.9), (6.12), and of matrix $\mathbf{F}^c$ in Section 1.2.3. Reformulation of $\rho\left(\mathbf{F}^c + \Delta\mathbf{F}^{c,box}\right)$ as follows, where $\|\cdot\|_F$ denotes the Frobenius norm, yields

$$\rho(\mathbf{F}^c + \Delta\mathbf{F}^{c,box}) \leq \|\mathbf{F}^c + \Delta\mathbf{F}^{c,box}\|_2 \leq \|\mathbf{F}^c\|_2 + \|\Delta\mathbf{F}^{c,box}\|_F \tag{A.37a}$$

$$= \|\mathbf{F}^c\|_2 + \Gamma\sqrt{\sum_{u \in \mathcal{U}}\gamma_u^c\sum_{\text{v} \in \mathcal{U}}|\frac{\bar{H}_{u\text{v}}^c}{H_{uu}^c}\Lambda_{u\text{v}}^c|^2} \tag{A.37b}$$

$$\leq \|\mathbf{F}^c\|_2 + \Gamma\sqrt{\sum_{u \in \mathcal{U}}\gamma_u^c\sum_{\text{v} \in \mathcal{U}}\left(\frac{\bar{H}_{u\text{v}}^c}{H_{uu}^c}\right)^2\varepsilon_u^2} \tag{A.37c}$$

$$= \|\mathbf{F}^c\|_2 + \Gamma\sqrt{\sum_{u \in \mathcal{U}}\varepsilon_u^2\gamma_u^c\sum_{\text{v} \in \mathcal{U}}\left(\frac{\bar{H}_{u\text{v}}^c}{H_{uu}^c}\right)^2} \tag{A.37d}$$

$$\leq \|\mathbf{F}^c\|_2 + \|\mathbf{F}^c\|_F \cdot \|\boldsymbol{\varepsilon}\|_2 \tag{A.37e}$$

$$\leq \|\mathbf{F}^c\|_2 \cdot \left(1 + \sqrt{U}\|\boldsymbol{\varepsilon}\|_2\right), \tag{A.37f}$$

where in particular (A.37b) follows from (A.36) and the definition of the Frobenius norm, and (A.37c) follows from the definition of $\Lambda^c_{uv}, \forall v \in \mathcal{U}$. Hence, $\|\mathbf{F}^c\|_2 \cdot \left(1 + \sqrt{U}\|\boldsymbol{\varepsilon}\|_2\right) < 1$ is sufficient for $\rho\left(\mathbf{F}^c + \Delta\mathbf{F}^{c,box}\right) < 1$ to hold. By (6.14) this condition is also sufficient for feasibility of the problem in (6.11), which proves the first part of the theorem. Having the same uncertainty parameter ε for all users we can write more simply $\Delta\mathbf{F}^{c,box} = \varepsilon\mathbf{F}^c$, and as the largest eigenvalue necessarily also scales with $(1 + \varepsilon)$ we obtain the second part of the theorem. $\qquad\square$

A.13 Proof of Theorem 14

The first part of the outage probability bounds can be immediately derived based on [10] as follows. The outage probability of an allocation $\mathbf{p}^c$ in Problem (6.13) under symmetric distribution of crosstalk coefficients in the interval given in (6.17) can (based on the constraint-inequalities in (6.13)) be written as

$$\sigma = Pr\{p^c_u - \Gamma\frac{\gamma^c_u}{H^c_{uu}}(\mathbf{p}^c)^T\bar{\mathbf{H}}^c_{u:} - \varepsilon_u\Gamma\frac{\gamma^c_u}{H^c_{uu}}(\mathbf{p}^c)^T \operatorname{diag}\left(\bar{\mathbf{H}}^c_{u:}\right)\boldsymbol{\eta}^u < n^c_u\}, \qquad (A.38)$$

where $\boldsymbol{\eta}^u \in \mathbb{R}^U$, $\eta^u_v \in [-1, 1]$ being independent, symmetrically distributed random variables following from the definition of the distribution in (6.17). A solution to (6.11) is necessarily feasible for the constraints in (6.11b), and we may hence insert these constraints into (A.38), giving

$$\sigma \leq Pr\{\varepsilon_u\Gamma\frac{\gamma^c_u}{H^c_{uu}}\|\operatorname{diag}(\bar{\mathbf{H}}^c_{u:})\mathbf{p}^c\|_2 + n^c_u - \varepsilon_u\Gamma\frac{\gamma^c_u}{H^c_{uu}}(\mathbf{p}^c)^T \operatorname{diag}\left(\bar{\mathbf{H}}^c_{u:}\right)\boldsymbol{\eta}^u < n^c_u\},$$
$$\qquad (A.39a)$$
$$= Pr\{(\mathbf{p}^c)^T \operatorname{diag}\left(\bar{\mathbf{H}}^c_{u:}\right)\boldsymbol{\eta}^u > \|\operatorname{diag}(\bar{\mathbf{H}}^c_{u:})\mathbf{p}^c\|_2\}. \qquad (A.39b)$$

The result follows together with the probability bound used in the proof to [10, Prop. 3.1]

$$Pr\left\{\sum_{v=1}^U a_v\eta_{uv} > \Omega\sqrt{\sum_{v=1}^U a_v^2}\right\} \leq e^{-\Omega^2/2}, \qquad (A.40)$$

where $a_v \in \mathbb{R}$, $\Omega \in \mathbb{R}_+$, η_{uv} are random variables with properties as specified above, and e is Euler's number.

In the case of uniformly distributed parameters $\tilde{H}^c_{uv}$ we may simply relate the volumes of ellipsoidal [63, p. 67] and box-shaped uncertainty regions from the equations in (6.7) and in (6.12) to obtain the second given outage probability bound. $\qquad\square$

Appendix B Algorithms

C.1 A Geometric Programming (GP) Approach for Line-Driver Power (LDP) Optimization

Geometric Programs (GPs) consist of posynomial objective and inequality constraints, as well as monomial equality constraints. Posynomial functions are sums $\sum_{k=1}^{K} f_k(\mathbf{p})$ of monomial functions $f_k(\mathbf{p}) : \mathbb{R}_{+}^{CU} \rightarrow \mathbb{R}$ of the form $f_k(\mathbf{p}) = c_k \cdot p_1^{\alpha_1^k} \cdot p_2^{\alpha_2^k} \cdot \ldots \cdot p_{UC}^{\alpha_{CU}^k}$, where $c_k \geq 0$ and $\alpha_i^k \in \mathbb{R}, 1 \leq i \leq CU$. Introducing auxiliary variables t_u, $u \in \mathcal{U}$, for the sum-power terms $\sum_{c \in \mathcal{C}} p_c^u$ in (2.24a), we obtain the equivalent formulation

$$\underset{p_c^u, t_u, u \in \mathcal{U}, c \in \mathcal{C}}{\text{minimize}} \quad \sum_{u \in \mathcal{U}} \sqrt{t_u} \tag{C.41a}$$

$$\text{subject to} \quad t_u^{-1} \cdot \sum_{c \in \mathcal{C}} p_c^u \leq 1, \forall u \in \mathcal{U}, \tag{C.41b}$$

$$\text{Constraints } (2.23b) - (2.23c). \tag{C.41c}$$

According to the definitions above, the objective in (C.41a) is a posynomial function and the auxiliary constraints in (C.41b) have posynomial form [21]. As noted in [38] the constraints in (1.8b) can also be written as posynomial constraints when using for instance the SINR approximation [130] $r_c^u(\mathbf{p}_c) \approx \tilde{r}_c^u(\mathbf{p}_c) = \alpha_c^u \log_2(\text{SINR}_c^u(\tilde{\mathbf{p}}_c)) + \beta_c^u, c \in \mathcal{C}, u \in \mathcal{U}$, where SINR_c^u is the SINR in (1.3) and $\tilde{p}_c^u, c \in \mathcal{C}, u \in \mathcal{U}$, is the approximation point. To see this, one needs to introduce additional variables $\tilde{t}_c^u$, $c \in \mathcal{C}, u \in \mathcal{U}$, which replace the total noise $(\sum_{i \in \mathcal{U} \setminus u} H_c^{ui} p_c^i + N_c^u)$ user u receives on subcarrier c. The thereby created additional constraints $1 \geq \frac{1}{\tilde{t}_c^u}(\sum_{i \in \mathcal{U} \setminus u} H_c^{ui} p_c^i + N_c^u)$, $c \in \mathcal{C}, u \in \mathcal{U}$, are posynomial expressions. Under these additional variables the constraints in (1.8c) and (2.23c) can be seen to be already given in posynomial and monomial form, respectively. Hence, we have that the problem in (C.41) can be approximated as a GP which can be efficiently solved optimally by convex optimization software [120].

Algorithm 10 Box-based Branch-and-Reduce Algorithm

1: Initialize the incumbent using a heuristic solution based
    on successive geometric programming, cf. Section 2.3.3.1.
2: Initialize the first open, currently active box with minimal
    and maximal corner-points $\mathbf{0} \in \mathbb{R}^{UC}$ and $\hat{\mathbf{p}} \in \mathbb{R}^{UC}$.
3: **while** {Any box is open} **do**
4:    <u>Branching:</u> Generate two new open boxes by splitting
       the currently active box in half in dB-scale in the
       dimension of its longest edge.*
5:    <u>Bounding:</u> Compute objective lower bounds for both
       new boxes using an underestimating LP [78] to the
       DCP problem in (2.33) with reduced variable ranges [142].
6:    <u>Reduction:</u> Try a range-reduction based on the current
       incumbent solution [142], and repeat the lower-bound
       LP if a range-reduction was achieved.
7:    <u>Incumbent Update:</u> Update the incumbent by testing the
       2^{CU-1} new corner points created through branching
       and the LP solutions for feasibility in (2.33).
8:    <u>Pruning:</u> Close all boxes with a lower bound above
       the incumbent solution.
9:    <u>Selection:</u> Choose the open box with the lowest
       lower bound as the new active box.
10: **end while**
    * In case the value of the minimal element in splitting dimension is zero we use a lower value
    based on a fixed ratio to the value of the maximal element in splitting dimension.

C.2 A Box-based Branch-and-Reduce Algorithm

Algorithm 10 schematically describes the proposed scheme for global optimization of the difference-of-convex-functions programming (DCP) problem in (2.33). The idea behind the method is to first enclose the set defined by the mask-constraints in (2.33c) by a box, cf. Line 2, and to successively split this set ("branching") into smaller boxes, cf. Line 4. We observed that box-based branching repeatedly outperforms simplicial branching [78]. We believe this is due to the conservative initial search space in simplicial branching, which is a simplex with corner points $\mathbf{0}, (\sum_{u \in \mathcal{U}, c \in \mathcal{C}} \hat{p}_c^u)\mathbf{e}_u, u \in \mathcal{U}$, where $\mathbf{e}_u$ is the u'th unit vector. Lower bounds on the objective value in any box are computed by linear programming (LP) after linearly approximating (underestimating) all convex functions $g_u(\mathbf{p})$ and all concave functions $-h_u(\mathbf{p}), u \in \mathcal{U}$, cf. Line 5. The fact that such a linear underestimation of convex and concave functions can easily be found [78] is the key advantage of the DCP formulation in (2.33). Differently to [78] we propose to apply linear approximations of all convex functions $g_u(\mathbf{p}), u \in \mathcal{U}$, not only on a single point but on various

points in the considered box, e.g., in regular intervals between the center point and each corner point. Based on the lower-bounds and the best feasible solution found so far (the "incumbent") the created boxes are either further split or discarded if the lower-bound lies above the upper bound, cf. Line 8. In [192] a transformation of variables into dB-scale was proposed. Similarly, we perform the branching (bisection) in dB-scale, which has the advantage that we still consider the full search-space beginning at a power-allocation of zero. Another technique integrated in Algorithm 10 is that of range reduction [142, 143]. Briefly speaking, bounds on the values that constraints may take in the LP used to compute lower bounds can be tightened based on the obtained optimal dual variables associated with these constraints and the current incumbent solution, cf. [142, 143] for details. Note that we omitted any local search step for improving the incumbent solution as is typically done in continuous BnB methods [143]. We believe the incumbent initialization in Line 1 by the successive geometric programming described in Appendix C.1 is tight enough for the considered applications to make such a local search in the BnB process redundant. We refer to [78] for a detailed description of a basic *simplicial* branch-and-bound algorithm applied to a general DCP problem, and to [142] for an introduction to the range-reduction technique, as well as to [26] for an application of range reduction in a specific DCP problem with DCP functions in the objective only.

Algorithm 11 DFB Tree Search

1: $[\mathbf{r}^*, f^*, f^{lb}] = \text{DFBSearch}(\mathbf{r}^0, \boldsymbol{\lambda}, \boldsymbol{\nu}, \Phi)$

....

2: ***Initialization—*** bEnd=false, $\mathbf{r} = \mathbf{0}$, $u = 1$, bRet=false, $\hat{\mathbf{r}} = \mathbf{1}\hat{\theta}$,
$\mathbf{p}^{\min} = \mathbf{0}$, $lb = -\infty$, $\hat{\mathbf{r}}^{\text{tmp}} = \hat{\mathbf{r}}$, $\mathbf{L}^{\text{tmp}} = \mathbf{L} = [\mathbf{0}, \hat{\mathbf{r}}]$, $f^{lb} = \infty$,

3: **if** $\mathbf{0} \preceq \mathbf{p}(\mathbf{r}^0) \preceq \hat{\mathbf{p}}$ and $f(\mathbf{p}(\mathbf{r}^0), \hat{\mathbf{w}} + \boldsymbol{\nu}, \breve{\mathbf{w}} + \boldsymbol{\lambda}) < 0$ **then**

4:     Set incumbent $\mathbf{r}^* = \mathbf{r}^0$, $f^* = f(\mathbf{p}(\mathbf{r}^0), \hat{\mathbf{w}} + \boldsymbol{\nu}, \breve{\mathbf{w}} + \boldsymbol{\lambda})$

5: **else** Set incumbent $\mathbf{r}^* = \mathbf{0}$, $f^* = 0$

6: $[\mathbf{L}_{u+1:U,:}, \mathbf{p}^{\min}, lb^{\text{SSR}}] = SSR(\boldsymbol{\lambda}, \boldsymbol{\nu}, \mathbf{r}, \hat{\mathbf{r}}^{\text{tmp}}, 0, f^*)$

7: **while** $\neg$bEnd

....

8:     Start search at $\mathbf{r} = \mathbf{L}_{:,1}$ — ***Update Allocation***

9:     **if** bRet **then** Set $r_u = r_u + \theta$ and bRet=false

10:         **if** $r_u > \min\{\hat{r}_u^{\text{tmp}}, \mathbf{L}_{u,2}^{\text{tmp}}\}$ **then** bRet=true   **else**

11:             **if** $\mathbf{0} \preceq \mathbf{p}(\mathbf{r}) \preceq \hat{\mathbf{p}}$ **then** Set $\mathbf{p}^{\min} = \mathbf{p}(\mathbf{r})$

12:                 **if** $f(\mathbf{p}(\mathbf{r}), \hat{\mathbf{w}} + \boldsymbol{\nu}, \breve{\mathbf{w}} + \boldsymbol{\lambda}) < f^*$ **then**

13:                     $\mathbf{r}^* = \mathbf{r}$, $f^* = f(\mathbf{p}(\mathbf{r}^*), \hat{\mathbf{w}} + \boldsymbol{\nu}, \breve{\mathbf{w}} + \boldsymbol{\lambda})$

14:             **else** bRet=true, $\hat{r}_u^{\text{tmp}} = r_u - \theta$

....

15:     **if** $\neg$bRet **then**   — ***Bounding***

16:         $[\mathbf{L}_{u+1:U,:}^{\text{tmp}}, \mathbf{p}^{\min}, lb^{\text{SSR}}] = SSR(\boldsymbol{\lambda}, \boldsymbol{\nu}, \mathbf{r}, \hat{\mathbf{r}}^{\text{tmp}}, u, f^*)$

17:         **if** $u < U$ **then**

$$r_i^{\max} = \begin{cases} r_i, & 1 \leq i \leq u, \\ \min\{\hat{r}_i^{\text{tmp}}, \mathbf{L}_{i,2}^{\text{tmp}}\}, & u + 1 \leq i \leq U \end{cases}$$

18:             $lb = lb^{\mathbf{r}^{(u)}}(\mathbf{p}^{\min}, \mathbf{r}^{\max})$, cf. (3.3), $lb = \max\{lb, lb^{\text{SSR}}\}$

....

19:             **if** $lb \geq \min\{f^*, \Phi\}$ **then**   — ***Branching***

20:                 Set bRet=true, $f^{lb} = \min\{f^{lb}, lb\}$

21:             **else** $u = u + 1$

22:     **else** Set $r_u = \mathbf{L}_{u,1}$, $u = u - 1$

23:         **if** $u < U$ **then** $\hat{r}_{u+1} = \mathbf{L}_{u+1,2}$

24:         **if** $u = 0$ **then** bEnd=true

25: **if** $f^* \leq \Phi$ **then** $f^{lb} = f^*$

Algorithm 12 Search-Space Reduction Scheme

1: $[\mathbf{L}_{u+1:U,:}^{\text{tmp}}, \mathbf{p}^{\min}, \check{L}] = SSR(\boldsymbol{\lambda}, \boldsymbol{\nu}, \mathbf{r}^{\text{init}}, \hat{\mathbf{r}}^{\text{tmp}}, u, f^*)$

2: Initialize $\mathbf{p} = \mathbf{p}(\mathbf{r}^{\text{init}})$, $\mathbf{p}^{\min} = \mathbf{p}$, $\bar{\mathbf{p}} = \mathbf{p}$, $\bar{\mathbf{r}} = \mathbf{r}^{\text{init}}$

$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$

3: ***Find the discrete optimum without crosstalk:***

4: Compute $\tilde{\mathbf{N}} \in \mathbb{R}^U$, $\tilde{N}_i = N_i + \sum_{1 \leq j \leq u} H_{ij} p_j, \forall i \in \mathcal{U}$

5: **Repeat** $\forall i \in \mathcal{U} \setminus \{1, \ldots, u\}$ Equations (A.22)-(A.24)

6: $\check{L} = (\hat{\mathbf{w}} + \boldsymbol{\nu})^T \bar{\mathbf{p}} - (\check{\mathbf{w}} + \boldsymbol{\lambda})^T \bar{\mathbf{r}}$, Set $\mathbf{r} = \bar{\mathbf{r}}$

7: **while** $(\mathbf{p}(\mathbf{r}) \notin \mathcal{Q})$ **do** $r_i = \max\{0, r_i - \theta\}, u < i \leq U$

8: $L^{\text{target}} = \min\{f^*, f(\mathbf{p}(\mathbf{r}), \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda})\}$

$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$

9: ***Find the restricted search-region around*** $\bar{\mathbf{r}}$***:***

10: **for** $i = u+1, \ldots, U$ **do**

11:     Set $\mathbf{r} = \bar{\mathbf{r}}$, **Increase** r_i in steps of θ **until** $r_i > \hat{r}_i^{\text{tmp}}$
        or $(\hat{\mathbf{w}} + \boldsymbol{\nu})^T \mathbf{p} - (\check{\mathbf{w}} + \boldsymbol{\lambda})^T \mathbf{r} > L^{\text{target}}$, with
        $p_i = p_i(\mathbf{r}^{\text{init}}), i \in [1, u], p_i = p_i^{\text{tmp}}(r_i)$ as in $(A.23), i \in \mathcal{U} \setminus [1, u]$,

12:     $\mathbf{L}_{i,2}^{\text{tmp}} = r_i - \theta$, $\mathbf{r} = \bar{\mathbf{r}}$

13:     **Decrease** r_i in steps of θ **until** $r_i < 0$ **or**
        $(\hat{\mathbf{w}} + \boldsymbol{\nu})^T \mathbf{p} - (\check{\mathbf{w}} + \boldsymbol{\lambda})^T \mathbf{r} > L^{\text{target}}$, with
        $p_i = p_i(\mathbf{r}^{\text{init}}), i \in [1, u], p_i = p_i^{\text{tmp}}(r_i)$ as in $(A.23), i \in \mathcal{U} \setminus [1, u]$,

14:     $\mathbf{L}_{i,1}^{\text{tmp}} = r_i + \theta$, $p_i^{\min} = p_i^{\text{tmp}}(r_i + \theta)$

15: **end for**

Algorithm 13 Grasp based Bit-Loading

1: Initialize $\mathbf{r} = \mathbf{0}$, $\mathbf{r}^*$, $\upsilon^* = 0$, $f^{\mathrm{prev}} = f(\mathbf{r})$, $\boldsymbol{\beta} \in \mathbb{R}^M$, $\beta_i \in [0,1]$, $g_i = \check{f}$, $\forall i \in \mathcal{M} = \{1, \ldots, M\}$, $k = 1$, K

2: **for** $k = 1, \ldots, K$ **do**

3:     **while** $\upsilon^{\mathrm{min}} \leq 0$ **do**

4:         Set $\beta = \beta_{i^*}$, $\tilde{g}_i = \max_{i \in \mathcal{M}}\{g_i\} - g_i, \forall i \in \mathcal{M}$, where $i^* \in \mathcal{M}$ is sampled from the distribution

$$P_i = \frac{\tilde{g}_i}{\sum_{j \in \mathcal{M}} \tilde{g}_j}, \ \forall i \in \mathcal{M}, \tag{C.42}$$

5:         **for** $u = 1, \ldots, U$ **do**

6:           **if** $\exists \mathbf{p} \in \mathcal{Q} | r_u(\mathbf{p}) = r_u + \theta, r_i(\mathbf{p}) = r_i, i \in \mathcal{U} \setminus \{u\}$,

7:           **then** $\upsilon_u = f(\mathbf{p}, \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda}) - f^{\mathrm{prev}}$

8:           **else** $\upsilon_u = \infty$

9:         **end for**

10:        $\upsilon^{\mathrm{min}} = \min_{u \in \mathcal{U}}\{\upsilon_u\}$, $\upsilon^{\mathrm{max}} = \max_{u \in \mathcal{U}}\{\upsilon_u\}$

11:        Uniformly sample a user u^* from the set $\{u \in \mathcal{U} \mid \upsilon_u \leq (1 - \beta)\upsilon^{\mathrm{min}} + \beta\upsilon^{\mathrm{max}}\}$

12:        **if** $\upsilon^{\mathrm{min}} \leq 0$ **then** $r_{u^*} = r_{u^*} + \theta$, $f^{\mathrm{prev}} = f^{\mathrm{prev}} + \upsilon_{u^*}$

13:     **end while**

14:     Search a local optimum $\tilde{\mathbf{r}}$ starting at $\mathbf{r}$, cf. Algorithm 5

15:     Update incumbent $\mathbf{r}^* = \tilde{\mathbf{r}}$ if $f(\tilde{\mathbf{r}}) < f(\mathbf{r}^*)$

16:     Update the average g_{i^*} with the newest value f^{prev}

17: **end for**

Algorithm 14 rSEGO

1: Initialize $\mathbf{r}^*$, K, $\bar{K}$, values $\tilde{\tau}_u(i)$, $\forall u, i \in \mathcal{U}$, q_0, β, ρ
2: **for** $k = 1, \ldots, K$ **do** {/∗iterations∗/}
3:     **for** $\bar{k} = 1, \ldots, \bar{K}$ **do** {/∗ant runs∗/}
4:        Set $\mathbf{r}^{(\bar{k})} = \mathbf{0}$, $\tilde{\mathcal{U}} = \mathcal{U}$, Uniformly sample $s_1^{(\bar{k})} \in \mathcal{U}$
5:        **for** $i = 1, \ldots, U$ **do** {/∗sequential decisions∗/}
6:           Set $u = s_i^{(\bar{k})}$, $f^{\min} = \min\limits_{\{r_u^{(\bar{k})} \in \mathcal{B} | \mathbf{p}(\mathbf{r}^{(\bar{k})}) \in \mathcal{Q}\}} \{f(\mathbf{r}^{(\bar{k})})\}$,
          $f^{\max} = \max_{\{r_u^{(\bar{k})} \in \mathcal{B} | \mathbf{p}(\mathbf{r}^{(\bar{k})}) \in \mathcal{Q}\}} \{f(\mathbf{r}^{(\bar{k})})\}$
7:           Uniformly sample $r_u^{(\bar{k})}$ from the set $\{r_u^{(\bar{k})} \in \mathcal{B} \mid$
          $\mathbf{p}(\mathbf{r}^{(\bar{k})}) \in \mathcal{Q}, f(\mathbf{r}^{(\bar{k})}) \leq (1 - \beta)f^{\min} + \beta f^{\max}\}$
8:           **if** $i < U$ **then**
9:              Exclude u from $\tilde{\mathcal{U}}$, Uniformly sample $q \in [0, 1]$
10:              **if** $q < q_0$ **then** $\{s_{i+1}^{(\bar{k})} = \operatorname{argmax}_{j \in \tilde{\mathcal{U}}}\{\tilde{\tau}_u(j)\}\}$
11:              **else**
12:                 Sample $s_{i+1}^{(\bar{k})} = j \in \mathcal{U}$ from $P_u(j \mid \tilde{\mathcal{U}}) = 0, j \in \mathcal{U} \setminus \tilde{\mathcal{U}}$, $P_u(j \mid \tilde{\mathcal{U}}) = $
                $\frac{\tilde{\tau}_u(j)}{\sum_{l \in \tilde{\mathcal{U}}} \tilde{\tau}_u(l)}, j \in \tilde{\mathcal{U}}$,
13:              **end if**
14:              Update $\tilde{\tau}_u(s_{i+1}^{(\bar{k})}) = (1 - \rho) \cdot \tilde{\tau}_u(s_{i+1}^{(\bar{k})})$
15:           **end if**
16:        **end for**
17:        Update $\mathbf{r}^{(\bar{k})}$ by local search starting at $\mathbf{r}^{(\bar{k})}$, cf. Algorithm 5
18:     **end for**
19:     $\bar{k}^* = \operatorname*{argmin}\limits_{\bar{k} = 1, \ldots, \bar{K}}\{f(\mathbf{r}^{(\bar{k})})\}$, Update $\tilde{\tau}_{s_i^{(\bar{k}^*)}}(s_{i+1}^{(\bar{k}^*)})$ as in (4.9), $\forall i \in \{1, \ldots, U - 1\}$, Set
    $\mathbf{r}^* = \mathbf{r}^{(\bar{k}^*)}$ if $f(\mathbf{r}^{(\bar{k}^*)}) < f(\mathbf{r}^*)$
20: **end for**

C.3 Heuristic for Joint Margin, Artificial Noise, and Transmit-Power Optimization

---

Algorithm 15 Low-Complexity DSM Heuristic for Joint SNR Margin, AN and Transmit-Power Optimization

---

1: *Optional Initialization:* Run Greedy Multi-user Bit-
    Loading [103] under the Initial SNR margin and Use a
    Heuristic$^{\otimes}$ to Recover a *Stable* (i.e., feasible w.r.t. (7.6c)
    Initial Bit- and Power-Allocation
2: **repeat**
3:    **for** $u = 1$ to U **do**
4:       **while** Line-Search not Finished **do**
5:          Update γ_u Inside a Line-Search*
6:          Modified$^{\otimes}$ Greedy Multi-user Bit-Loading [103]
               - Initialization based on Incumbent Allocation
               - Calculate the Cost in (1.9) by Solving (7.6)
7:       **end while**
8:    **end for**
9: **until** Incumbent Objective Change is Below a Threshold
10: *Optional:* Call Algorithm 8 for the Found Incumbent SNR Margins to Obtain
    the Final Bit, Power, and AN Allocation

$^{\otimes}$ The incumbent power-allocation (as currently available for any SNR margins in Line 6 or in Line 1 the unstable power-allocation obtained by the algorithm in [103]) is used to produce an initial discrete bit-allocation, which is decreased by θ for all users until it is feasible in (7.6) under the actual SNR margins.

* Initially we halve Δ until $\gamma_u + \Delta$ or $\gamma_u - \Delta$ gives an improving objective ("Local Search"), determining a search direction $d = +1$ or $d = -1$, respectively, and an update $\gamma_u = \gamma_u + d \cdot \Delta$. Next we continue by searching in the improving direction ("directed search"), i.e., we repeat $\gamma_u = \gamma_u + d \cdot \Delta$ while an improvement in objective is made. Otherwise we halve Δ and go back to the "local search" step. The process finishes when the step-size Δ is below a threshold $\Delta^{\min}$ and hence is guaranteed to converge.

Algorithm 15 summarizes a scheme for the joint optimization of the SNR margin and the spectral AN (or VN) and transmit-power allocation. Similarly as in [119] we apply a margin search on top of a common DSM algorithm. However, in [119] a margin maximization was targeted, while differently our objective is the optimization of power and/or rate under stability constraints as defined in Section 7.1. In Line 6 we run a greedy multi-user discrete bit-loading algorithm, differing from that in [103] solely by the initialization which considers the currently best found margin, AN, and transmit-power allocation, and the procedure for computing the cost in (1.9) of a specific bit-allocation, namely by solving the LP in (7.6) (or the problem in (7.7)

for the "margin-only" case, or a similar problem as in (7.6) in case of using VN for $\gamma_u > 1$ and $v_u^c \geq 0$, cf. Sections 7.1.3–7.1.5 for details). In Lines 4–6 we perform a line-search for the parameter γ_u of user u, which is repeated for all users. As commented in Line 10 one may choose to call in the end a more complex DSM algorithm for the found incumbent/fixed SNR margins as in Algorithm 8 or a similar scheme using suboptimal heuristics, cf. the discussion of alternatives in Section 7.2.

Bibliography

[1] S.-J. Kim A. Mutapcic and S. Boyd. Beamforming with uncertain weights. *IEEE Signal Processing Letters*, 14(5):348–351, May 2007.

[2] K.M. Abadir and J.R. Magnus. *Matrix Algebra*. Cambridge University Press, 2005.

[3] C.H. Aldana and J.M. Cioffi. Channel tracking for multiple input, single output systems using EM algorithm. In *IEEE International Conference on Communications 2001 (ICC'01)*, volume 2, pages 586–590, Helsinki, Finland, 11-14 June 2001.

[4] J. P. Aubin and I. Ekeland. Estimates of the duality gap in nonconvex optimization. *Mathematics of Operations Research*, 1(3):225–245, August 1976.

[5] E. Baccarelli and M. Biagi. Optimal integer bit-loading for multicarrier ADSL systems subject to spectral-compatibility limits. *Signal Processing*, 84(4):729–741, April 2004.

[6] D.E. Barreto and S.S. Chiu. Decomposition methods for cross-layer optimization in wireless networks. In *IEEE Wireless Communications and Networking Conference*, pages 270–275, Hong Kong, China, 11-15 March 2007.

[7] M. Barton and M. Honig. Spectral optimization of discrete multitone system on twisted wire copper plants. In *IEEE International Conference on Communications 1995 (ICC'95)*, volume 3, pages 1668–1672, Seattle, USA, 18-22 June 1995.

[8] P. Belotti. Couenne: A user's manual. Technical report, Lehigh University, 2009.

[9] P. Belotti, J. Lee, L. Liberti, F. Margot, and A. Wächter. Branching and bounds tightening techniques for non-convex MINLP. *Optimization Methods and Software*, 24(4-5):597–634, 2009.

[10] A. Ben-Tal and A. Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming*, 88:411–424, 2000.

[11] D. P. Bertsekas. *Constrained optimization and Lagrange multiplier methods.* Academic Press, 1982.

[12] D. P. Bertsekas. *Nonlinear Programming.* Athena Scientific, 1999.

[13] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.

[14] D. Bertsimas and J. N. Tsitsiklis. *Introduction to linear optimization.* Athena Scientific, 1997.

[15] S. Bhaumik, D. Chuck, G. Narlikar, and G. Wilfong. Energy-efficient design and optimization of wireline access networks. In *IEEE International Conference on Computer Communications 2011 (INFOCOM'11)*, pages 451–455, Shanghai, China, 10–15 April 2011.

[16] C. Bianco, F. Cucchietti, G. Griffa, Kong Xiaoming, Chen Qiao, Hong Yuping, P. Gemma, and Zhai Liqian. An update on the field trial concerning underground solution for FTTCab architecture. In *IEEE International Telecommunications Energy Conference 2009 (INTELEC'09)*, Incheon, South Korea, 18-22 October 2009.

[17] C. Bianco, F. Cucchietti, G. Griffa, Hong Yuping, Kong Xiaoming, Chen Qiao, Luo Shudong, and Han Dong. An update on the field trial concerning free cooling solution for FTTCab architecture. In *IEEE International Telecommunications Energy Conference 2009 (INTELEC'09)*, Incheon, South Korea, 18-22 October 2009.

[18] C. Bianco, G. Griffa, N. Lee, P. Gemma, and Zhu Bin. The energy saving evaluation of Green DSL while in-field mass implementation. In *IEEE International Telecommunications Energy Conference 2011 (INTELEC'11)*, Amsterdam, The Netherlands, 9-13 October 2011.

[19] P. Björklund, P. Värbrand, and D. Yuan. Resource optimization of spatial TDMA in ad hoc radio networks: A column generation approach. In *IEEE International Conference on Computer Communications 2003 (INFOCOM'03)*, volume 2, pages 818–824, San Francisco, CA, USA, 30 March – 3 April 2003.

[20] R. Bolla, F. Davoli, R. Bruschi, K. Christensen, F. Cucchietti, and S. Singh. The potential impact of green technologies in next-generation wireline networks: Is there room for energy saving optimization? *IEEE Communications Magazine*, 49(8):80–86, August 2011.

[21] S. Boyd, S.-J. Kim, L. Vandenberghe, and A. Hassibi. A tutorial on geometric programming. *Optimization and Engineering*, 8(1):67–127, March 2007.

[22] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[23] J. Brehmer, Qing Bai, and W. Utschick. Time-sharing solutions in MIMO broadcast channel utility maximization. In *International ITG Workshop on Smart Antennas 2008 (WSA'08)*, pages 153–156, Darmstadt, Germany, 26-27 February 2008.

[24] J. Brehmer and W. Utschick. A decomposition of the downlink utiliy maximization problem. In *Conference on Information Sciences and Systems (CISS'07)*, pages 437–441, Baltimore, MD, USA, 14–16 March 2007.

[25] J. Brehmer and W. Utschick. Nonconcave utility maximization in the OFDMA downlink. In *International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks and Workshops (WiOpt'08)*, pages 413–418, Berlin, Germany, 31 March - 4 April 2008.

[26] R. Cambini and F. Salvi. A branch and reduce approach for solving a class of low rank d.c. programs. *Journal of Computational and Applied Mathematics*, 233(2):492–501, November 2009.

[27] J. Campello. Optimal discrete bit loading for multicarrier modulation systems. In *IEEE International Symposium on Information Theory 1998 (ISIT'98)*, page 193, Cambridge, MA, USA, 16-21 August 1998.

[28] R. Cendrillon, G. Ginis, E. Van den Bogaert, and M. Moonen. A near-optimal linear crosstalk canceler for upstream VDSL. *IEEE Transactions on Signal Processing*, 54(8):3136–3146, August 2006.

[29] R. Cendrillon, Jianwei Huang, Mung Chiang, and M. Moonen. Autonomous spectrum balancing for digital subscriber lines. *IEEE Transactions on Signal Processing*, 55(8):4241–4257, August 2007.

[30] R. Cendrillon, Fang Liming, J. Chou, Guozhu Long, Chin Hung, and Dong Wei. DSM from theory to practice. In *IEEE Global Communications Conference 2008 (Globecom'08)*, New Orleans, LA, USA, 30 November – 4 December 2008.

[31] R. Cendrillon and M. Moonen. Iterative spectrum balancing for digital subscriber lines. In *IEEE International Conference on Communications 2005 (ICC'05)*, volume 3, pages 1937–1941, Seoul, Korea, 16-20 May 2005.

[32] R. Cendrillon, M. Moonen, D. Gore, and A. Paulraj. Low complexity crosstalk cancellation through line selection in upstream VDSL. In *International Conference on Acoustics, Speech and Signal Processing 2003 (ICASSP'03)*, volume 4, pages 692–695, Hong Kong, China, 6–10 April 2003.

[33] R. Cendrillon, Wei Yu, M. Moonen, J. Verlinden, and T. Bostoen. Optimal multiuser spectrum balancing for digital subscriber lines. *IEEE Transactions on Communications*, 54(5):922–933, May 2006.

[34] V.M.K. Chan and Wei Yu. Multiuser spectrum optimization for discrete multitone systems with asynchronous crosstalk. *IEEE Transactions on Signal Processing*, 55(11):5425–5435, November 2007.

[35] M. Charafeddine and A. Paulraj. Sequential geometric programming for 2×2 interference channel power control. In *Conference on Information Sciences and Systems (CISS'07)*, pages 185–189, Baltimore, MD, USA, 14-16 March 2007.

[36] M. Chatterjee, H. Lin, and S.K. Das. Rate allocation and admission control for differentiated services in CDMA data networks. *IEEE Transactions on Mobile Computing*, 6(2):179–191, February 2007.

[37] M. Chiang, P. Hande, T. Lan, and C.W. Tan. Power control in wireless cellular networks. In *Foundations and Trends in Networking Sample*, volume 2 (4), pages 381–533. Now Publishers, July 2008.

[38] M. Chiang, C.W. Tan, D.P. Palomar, D. O'Neill, and D. Julian. Power control by geometric programming. *IEEE Transactions on Wireless Communications*, 6(7):2640 – 2651, July 2007.

[39] Mung Chiang. *Geometric Programming for Communication Systems*. Now Publishers Inc, 2005.

[40] J.M. Cioffi and M. Mohseni. Dynamic spectrum management. *Telektronikk*, 4:126–137, 2004.

[41] J.M. Cioffi, H. Zou, A. Chowdhery, W. Lee, and S. Jagannathan. Greener copper with dynamic spectrum management. In *IEEE Global Telecommunications Conference 2008 (Globecom'08)*, New Orleans, LA, USA, 30 November – 4 December 2008.

[42] Network Interoperability Consultative Committee. Guidelines on DSL power saving modes and non-stationary noise in metallic access networks. Technical Report ND1424:2008/02, v. 1.1.1, NICC, 2008.

[43] T.M. Cover and J.A. Thomas. *Elements of Information Theory*. Wiley Interscience, Hoboken, NJ, USA, 2nd edition, 2006.

[44] G.B. Dantzig. *Linear programming and extensions*. Princeton University Press, 1963.

[45] E.D. Dolan and J.J. Moré. Benchmarking optimization software with performance profiles. *Mathematical Programming*, 91(2):201–213, 2002.

[46] M. Dorigo and L.M. Gambardella. Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation*, 1(1):53–66, April 1997.

[47] P.E. Eriksson E. Trojer. Power saving modes for GPON and VDSL2. `http://www.ieee802.org/3/10GEPON_study/email/pdfV3kikUObAl.pdf`, 2008.

[48] EC Directorate-General JRC Joint Research Centre, Institute for Energy, Renewable Energy Unit. Code of conduct on energy consumption of broadband equipment. Version 4, February 2011.

[49] K. Eriksson, S. Shi, N. Vucic, M. Schubert, and E.G. Larsson. Globally optimal resource allocation for achieving maximum weighted sum rate. In *IEEE Global Telecommunications Conference (Globecom'10)*, Miami, USA, 6–10 December 2010.

[50] ETSI. Transmission and Multiplexing (TM); Access transmission systems on metallic access cables; Very high speed Digital Subscriber Line (VDSL); Part 1: Functional requirements. TM6 TS 101 270-1, Version 1.3.1, July 2003.

[51] ETSI. Environmental engineering (EE); the reduction of energy consumption in telecommunications equipment and related infrastructure. ETSI TS 102 530, Version 1.2.1, July 2011.

[52] A.R. Forouzan. Optimal spectrum management of DSL with nonstrictly convex rate region. *IEEE Transactions on Signal Processing*, 57(7):2558–2568, July 2009.

[53] A.R. Forouzan and M. Moonen. Lagrange multiplier optimization for optimal spectrum balancing of DSL with logarithmic complexity. In *IEEE International Conference on Communications 2011 (ICC'11)*, Kyoto, Japan, 5-9 June 2011.

[54] G. J. Foschini and Z. Miljanic. A simple distributed autonomous power control algorithm and its convergence. *IEEE Transactions on Vehicular Technology*, 42(4):641–646, November 1993.

[55] S. Galli, C. Valenti, and K.J. Kerpez. A frequency-domain approach to crosstalk identification in xDSL systems. *IEEE Journal of Selected Areas in Communications*, 19(8):1497–1506, August 2001.

[56] A. Garg and A.K. Chaturvedi. Generalization of hybrid time divisioning for power allocation in DMT-based DSL systems. *IEEE Communications Letters*, 11(6):504–506, June 2007.

[57] A. Ghosh and S. Boyd. Minimax and convex-concave games. Tutorial for course EE392o, Autumn 2003. Stanford University.

[58] G. Ginis. Low-power modes for ADSL2 and ADSL2+. Technical report, Broadband Communications Group, Texas Instruments, January 2005.

[59] P. Golden, H. Dedieu, and K.S. Jacobsen, editors. *Fundamentals of DSL technology.* Auerbach Publications, 2006.

[60] P. Golden, H. Dedieu, and K.S. Jacobsen, editors. *Implementation and Application of DSL Technology.* Auerbach Publications, 2008.

[61] E. Goma, M. Canini, A.L. Toledo, N. Laoutaris, D. Kostić, P. Rodriguez, R. Stanojević, and P.Y. Valentín. Insomnia in the access. In *ACM SIGCOMM 2011*, Toronto, Canada, 15-19 August 2011.

[62] G. Griffa, L. Radice, and C. Bianco. Carbon footprint of next generation fixed networks. In *IEEE International Telecommunications Energy Conference 2010 (INTELEC'10)*, Orlando, Florida, USA, 6-10 June 2010.

[63] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization.* Springer, 1988.

[64] The Climate Group. SMART 2020: Enabling the low carbon economy in the information age, 2008.

[65] M. Guenach, C. Nuzman, K. Hooghe, J. Maes, and M. Peeters. On power-efficient usage of line drivers in copper-based access networks. In *IEEE International Energy Conference and Exhibition (EnergyCon'10)*, pages 131–136, Manama, Bahrain, 18-22 December 2010.

[66] M. Guenach, C. Nuzman, K. Hooghe, J. Maes, and M. Peeters. Reduced dimensional power optimization using class AB and G line drivers in DSL. In *International Workshop on Green Communications, IEEE Globecom 2010*, pages 1443–1447, Miami, USA, 6–10 December 2010.

[67] M. Guenach, C. Nuzman, J. Maes, and M. Peeters. On power optimization in DSL systems. In *IEEE International Conference on Communications Workshops (ICC'09)*, Dresden, Germany, 14-18 June 2009.

[68] M. Guenach, C. Nuzman, J. Maes, and M. Peeters. Trading off rate and power consumption in DSL systems. In *International Workshop on Green Communications, IEEE Globecom 2009*, Honolulu, Hawaii, 30 November–4 December 2009.

[69] M. Guenach, C. Nuzman, J. Maes, M. Peeters, Y. Li, D. Van Bruyssel, and F. Defoort. Power efficient copper access. *Bell Labs Technical Journal*, 15(2):117–129, September 2010.

[70] E. Guerrini and D. Veronesi. Bit loading algorithm based on a probabilistic approach for HomePlug AV. In *IEEE Global Communications Conference 2011 (Globecom'11)*, Houston, Texas, USA, 5-9 December 2011.

[71] P. Hande, S. Zhang, and M. Chiang. Distributed rate allocation for inelastic flows. *IEEE Transactions on Networking*, 15(6):1240–1253, December 2007.

[72] S. Hayashi and Z.-Q. Luo. Spectrum management for interference-limited multiuser communication systems. *IEEE Transactions on Information Theory*, 55(3):1153–1175, March 2009.

[73] L.M.C. Hoo, B. Halder, J. Tellado, and J.M. Cioffi. Multiuser transmit optimization for multicarrier broadcast channels: Asymptotic FDMA capacity region and algorithms. *IEEE Transactions on Communications*, 52(6):922–930, June 2004.

[74] K. Hooghe and M. Guenach. Impact of FTTN architecture on access node energy efficiency. In *IEEE Symposium on Communications and Vehicular Technology in the Benelux 2010 (SCVT'10)*, Enschede, The Netherlands, 24-25 November 2010.

[75] K. Hooghe and M. Guenach. Toward green copper broadband access networks. *IEEE Communications Magazine*, 49(8):87–93, August 2011.

[76] K. Hooghe and M. Guenach. Towards energy-efficient packet processing in access nodes. In *IEEE Global Communications Conference 2011 (Globecom'11)*, Houston, Texas, USA, 5-9 December 2011.

[77] R. Horst and P.M. Pardalos, editors. *Handbook of Global Optimization*, volume 2 of *Nonconvex Optimization and its Applications*. Kluwer Academic Publishers, 1995.

[78] R. Horst and N.V. Thoai. DC programming: Overview. *Journal of Optimization Theory and Applications*, 103(1):1–43, October 1999.

[79] J. Huang, V.G. Subramanian, R. Agrawal, and R. Berry. Joint scheduling and resource allocation in uplink OFDM systems for broadband wireless access networks. *IEEE Journal on Selected Areas in Communications*, 27(2):226–234, February 2009.

[80] J. Huang, V.G. Subramanian, R. Agrawal, and R.A. Berry. Downlink scheduling and resource allocation for OFDM systems. *IEEE Transactions on Wireless Communications*, 8(1):288–296, January 2009.

[81] Y. Huang, R.H. Gohary, and Z.-Q. Luo. Approaching user capacity in a DSL system via harmonic mean-rate optimization. In *IEEE International Conference on Acoustics, Speech and Signal Processing 2009 (ICASSP'09)*, pages 2365 – 2368, Taipei, Taiwan, China, 19-24 April 2009.

[82] S. Huberman, C. Leung, and T. Le-Ngoc. Dynamic spectrum management (DSM) algorithms for multi-user xDSL. *IEEE Communications Surveys & Tutorials*, 14(1), First Quarter 2012.

[83] L.D. Humphrey. ADSL2/ADSL2plus low-power mode guidelines. Technical Report TR-202, Issue 1, The Broadband Forum, February 2010.

[84] IEEE. P802.3az energy efficient Ethernet task force, November 2007.

[85] Home Gateway Initiative. Requirements for an energy efficient home gateway. Technical Report HGI-RD009-R3, October 2010.

[86] ITU-T. Asymmetric digital subscriber line transceivers 2 (ADSL2). G.992.3, January 2005.

[87] ITU-T. Asymmetric digital subscriber line transceivers (ADSL) - Extended bandwidth ADSL2 (ADSL2+). G.992.5, January 2005.

[88] ITU-T. Very high speed digital subscriber line transceivers 2 (VDSL2). G.993.2, February 2006.

[89] ITU-T Study Group 15. Power saving tutorials. Tutorials and Presentations, Geneva, Switzerland, 13–15 February 2008.

[90] S. Jagannathan, Ch.S. Hwang, and J.M. Cioffi. Margin optimization in digital subscriber lines employing level-2 dynamic spectrum management. In *IEEE International Conference on Communications 2008 (ICC'08)*, pages 435–440, Beijing, China, 19–23 May 2008.

[91] M. Johansson and L. Xiao. Cross-layer optimization of wireless networks using nonlinear column generation. *IEEE Transactions on Wireless Communications*, 5(2):435–445, February 2006.

[92] A. Kalakech. *DSL: Crosstalk Estimation Techniques, and Impact on DSM.* PhD thesis, Université catholique de Lovain, 1348 Louvain-la-Neuve, Belgium, September 2010.

[93] A. Kalakech, J. Louveaux, and L. Vandendorpe. Applying the balanced capacity concept to DSL systems. In *IEEE International Conference on Communications 2007 (ICC'07)*, pages 2717–2721, Glasgow, Scotland, 24-28 June 2007.

[94] P. Kall and S.W. Wallace. *Stochastic programming.* John Wiley & Sons, 2nd edition, 1994.

[95] I. Kamitsos, P. Tsiaflakis, Sangtae Ha, and Mung Chiang. Stable sleeping in DSL broadband access: Feasibility and tradeoffs. In *IEEE Global Communications Conference 2011 (Globecom'11)*, Houston, Texas, USA, 5-9 December 2011.

[96] S. Kandukuri and S. Boyd. Optimal power control in interference limited fading wireless channels with outage probability specifications. *IEEE Journal on Wireless Communications*, 1(1):46–55, January 2002.

[97] K.J. Kerpez and S. Galli. Single-ended loop-makeup identification – Part II: Improved algorithms and performance results. *IEEE Transactions on Instrumentation and Measurement*, 55(2):538–549, April 2006.

[98] K.J. Kerpez, D.L. Waring, S. Galli, J. Dixon, and P. Madon. Advanced DSL management. *IEEE Communications Magazine*, 41(9):116–123, September 2003.

[99] J. Kors. Press release of the Bavarian regulatory authority for commercial broadcasting. `http://www.blm.de/de/pub/aktuelles/pressemitteilungen.cfm?eventPress=press.DisplayDetail&pressrelease_ID=1644`, July 2011.

[100] B.S. Krongold. Optimal efficient discrete resource allocation for OFDMA systems. In *IEEE Wireless Communications and Networking Conference 2008 (WCNC'08)*, pages 1722–1727, Las Vegas, Nevada, USA, 31 March - 3 April 2008.

[101] A. H. Land and A. G. Doig. An automatic method of solving discrete programming problems. *Econometrica*, 28(3):497–520, July 1960.

[102] M. Laurent and F. Rendl. Semidefinite programming and integer programming. In K. Aardal, G.L. Nemhauser, and R. Weismantel, editors, *Discrete Optimization*, volume 12 of *Handbooks in Operations Research and Management Science*, chapter 8, pages 393–514. Elsevier B.V., December 2005.

[103] J. Lee, R.V. Sonalkar, and J.M. Cioffi. Multi-user discrete bit-loading for DMT-based DSL systems. In *IEEE Global Telecommunications Conference 2002 (Globecom'02)*, volume 2, pages 1259–1263, Taipei, Taiwan, China, 17–21 November 2002.

[104] Ch. Leung, S. Huberman, and Tho Le-Ngoc. Autonomous spectrum balancing using multiple reference lines for digital subscriber lines. In *IEEE Global Communications Conference 2010 (Globecom'10)*, Miami, Florida, USA, 6-10 December 2010.

[105] K.K. Leung and L.-Ch. Wang. Integrated link adaptation and power control to improve error and throughput performance in broadband wireless packet networks. *IEEE Transactions on Wireless Communications*, 1(4):619–629, October 2002.

[106] F. Lindqvist, N. Lindqvist, and B. Dortschy et al. Crosstalk channel estimation via standardized two-port measurements. *EURASIP Journal on Advances in Signal Processing*, 2008, December 2008.

[107] N. Lindqvist, F. Lindqvist, M. Monteiro, B. Dortschy, E. Pelaes, and A. Klautau. Impact of crosstalk channel estimation on the DSM performance for DSL networks. *EURASIP Journal on Advances in Signal Processing*, 2010, February 2010.

[108] M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, 284:193–228, November 1998.

[109] R. Lorenz and S. Boyd. Robust minimum variance beamforming. *IEEE Transactions on Signal Processing*, 53(5):1684–1696, May 2005.

[110] M.E. Lübbecke and J. Desrosiers. Selected topics in column generation. *Operations Research*, 53(6):1007–1023, November-December 2005.

[111] R. Lui and W. Yu. Low-complexity near-optimal spectrum balancing for digital subscriber lines. In *IEEE International Conference on Communications 2005 (ICC'05)*, volume 3, pages 1947–1951, Seoul, Korea, 16-20 May 2005.

[112] Z.-Q. Luo and W. Yu. An introduction to convex optimization for communications and signal processing. *IEEE Journal of Selected Areas in Communications*, 24(8):1426–1438, August 2006.

[113] Z.-Q. Luo and S. Zhang. Dynamic spectrum management: Complexity and duality. *IEEE Journal of Selected Topics in Signal Processing*, 2(1):57–73, February 2008.

[114] Z.-Q. Luo and S. Zhang. Duality gap estimation and polynomial time approximation for optimal spectrum management. *IEEE Transactions on Signal Processing*, 57(7):2675 – 2689, July 2009.

[115] J. Maes, M. Guenach, and M. Peeters. Statistical channel model for gain quantification of DSL crosstalk mitigation techniques. In *IEEE International Conference on Communications 2009 (ICC'09)*, Dresden, Germany, 14-18 June 2009.

[116] J. Maes, M. Peeters, M. Guenach, and Ch. Storry. Maximizing digital subscriber line performance. *Bell Labs Technical Journal*, 13(1):105–116, Spring 2008.

[117] E. Medeiros, N. Lindqvist, M. Monteiro, H. Abrahim, F. Lindqvist, B. Dortschy, and A. Klautau. DSM performance on practical DSL systems based on estimated crosstalk channel information. In *European Signal Processing Conference 2009 (Eusipco'09)*, pages 2092–2096, Glasgow, Scotland, 24-28 August 2009.

[118] Farhad Meshkati, H.V. Poor, and S.C. Schwartz. Energy-efficient resource allocation in wireless networks. *IEEE Signal Processing Magazine*, pages 58–68, May 2007.

[119] M. Monteiro, A. Gomes, N. Lindqvist, B. Dortschy, and A. Klautau. An algorithm for improved stability of DSL networks using spectrum balancing. In *IEEE Global Communications Conference 2010 (Globecom'10)*, Miami, Florida, USA, 6-10 December 2010.

[120] MOSEK ApS, Denmark. *The MOSEK optimization toolbox for MATLAB manual, Version 6.0 (Revision 106)*, 2011.

[121] A. Mutapcic, S.-J. Kim, and S. Boyd. Array signal processing with robust rejection constraints via second-order cone programming. In *Asilomar Conference on Signals, Systems and Computers, 2006 (ACSSC '06)*, pages 2267–2270, Pacific Grove, CA, USA, 29 October – 1 November 2006.

[122] A. Mutapcic, S.-J. Kim, and S. Boyd. Robust Chebyshev FIR equalization. In *IEEE Global Telecommunications Conference 2007 (Globecom'07)*, pages 3074 – 3079, Washington, DC, USA, 26–30 November 2007.

[123] Y. Nakagawa. An improved surrogate constraints method for separable nonlinear integer programming. *Journal of Operations Research*, 46(2):145–163, 2003.

[124] G. L. Nemhauser and L. A. Wolsey. *Integer and Combinatorial Optimization*. Wiley Interscience, 1988.

[125] NIPP-NAI. Dynamic spectrum management technical report. Technical Report Pre-published document ATIS-PP-0600007, ATIS, May 2007.

[126] T. Nordström et.al. xDSL simulator v3.1. xdsl.ftw.at, October 2008.

[127] ATIS Exploratory Group on Green. Report on environmental sustainability, March 2009.

[128] N. Papandreou and T. Antonakopoulos. Far-end crosstalk identification method based on channel training sequences. *IEEE Transactions on Instrumentation and Measurement*, 54(6):2204–2212, December 2005.

[129] J. Papandriopoulos and J. S. Evans. Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks. In *IEEE International Conference on Communications 2006 (ICC'06)*, volume 7, pages 3270–3275, Istanbul, Turkey, 11-15 June 2006.

[130] J. Papandriopoulos and J.S. Evans. SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks. *IEEE Transactions on Information Theory*, 55(8):3711–3724, August 2009.

[131] A. Pascual-Iserte, D. Palomar, A. Perez-Neira, and M. Lagunas. A robust maximin approach for MIMO communications with imperfect channel state information based on convex optimization. *IEEE Transactions on Signal Processing*, 54:346–360, January 2006.

[132] M. Patriksson. Simplicial decomposition algorithms. In Ch.A. Floudas and P.M. Pardalos, editors, *Encyclopedia of Optimization*, pages 3579–3585. Springer, 2nd edition, 2009.

[133] J. Pierdomenico, S. Wurcer, and B. Day. A 684-mW adaptive supply full-rate ADSL CO driver. *IEEE Journal of Solid-State Circuits*, 37(12):1831–1838, December 2002.

[134] T. Piessens and M. Steyaert. *Design and analysis of high efficiency line drivers for xDSL*. Kluwer Academic Publishers, 2004.

[135] M. Pischella and J.-C. Belfiore. Weighted sum throughput maximization in multicell OFDMA networks. *IEEE Transactions on Vehicular Technology*, 59(2):896–905, February 2010.

[136] Europe's Energy Portal. Europe's energy portal. www.energy.eu, August 2011.

[137] B.D. Putra, T. Nordström, and S. Trautmann. Modeling Energy Efficiency in DSL Systems. Technical Report FTW-TR-2011-001, The Telecommunications Research Center Vienna (FTW), Vienna, Austria, May 2011.

[138] Li Ping Qian, Ying Jun Zhang, and Jianwei Huang. MAPEL: Achieving global optimality for a non-convex wireless power control problem. *IEEE Transactions on Wireless Communications*, 8(3):1536–1563, March 2009.

[139] M.G.C. Resende and C.C. Ribeiro. Greedy randomized adaptive search procedures. In F. Glover and G.A. Kochenberger, editors, *Handbook of Metaheuristics*, chapter 8, pages 219–249. Kluwer Academic Publishers, 2003.

[140] W. Rhee and J.M. Cioffi. Increase in capacity of multiuser OFDM system using dynamic subchannel allocation. In *IEEE Vehicular Technology Conference 2000 (VTC'00)*, volume 2, pages 1085–1089, Tokyo, Japan, 15–18 May 2000.

[141] S.N. Roy. Energy logic: A road map to reducing energy consumption in telecommunications networks. In *International Telecommunications Energy Conference 2008 (INTELEC'08)*, San Diego, CA, USA, 14-18 September 2008.

[142] H.S. Ryoo and N.V. Sahinidis. A branch-and-reduce approach to global optimization. *Journal of Global Optimization*, 8(2):107–138, 1996.

[143] N.V. Sahinidis. *BARON - Branch and Reduce Optimization Navigator*. University of Illinois at Urbana-Champaign, version 4.0 edition, June 2000.

[144] Sandvine. Global Internet phenomena report fall 2011, 2011.

[145] T. Sartenaer, L. Vandendorpe, and J. Louveaux. Balanced capacity of wireline multiuser channels. *IEEE Transactions on Communications*, 53(12):2029–2042, December 2005.

[146] K. Seong, M. Mohseni, and J.M. Cioffi. Optimal resource allocation for OFDMA downlink systems. In *IEEE International Symposium on Information Theory 2006 (ISIT'06)*, pages 1394–1398, Seattle, USA, 9-14 July 2006.

[147] B. Serneels and M. Steyaert. *Design of High Voltage xDSL Line Drivers in Standard CMOS*. Springer, 2008.

[148] P. Soldati and M. Johansson. Dynamic resource allocation in OFDMA multicellular systems. Technical Report TRITA-EE 2008:048, School of Electrical Engineering (KTH), August 2008.

[149] Kee Bong Song, Seong Taek Chung, G. Ginis, and J.M. Cioffi. Dynamic spectrum management for next-generation DSL systems. *IEEE Communications Magazine*, 40(10):101–109, October 2002.

[150] 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*, Princeton, NJ, USA, 30 April - 2 May 2007.

[151] R. Srikant. *The Mathematics of Internet Congestion Control.* Birkhauser, 2004.

[152] T. Starr, M. Sorbara, J.M. Cioffi, and P.J. Silverman. *DSL Advances.* PTR Communications Engineering and Emerging Technologies Series. Prentice Hall, 2003.

[153] D. Statovci and T. Nordström. Adaptive resource allocation in multiuser FDD-DMT systems. In *European Signal Processing Conference 2004 (Eusipco'04)*, pages 1213–1216, Vienna, Austria, 7-10 September 2004.

[154] D. Statovci and T. Nordström. Performance evaluation of the cable bundle unique power back-off algorithm. In *International Conference on Access Networks (AccessNets'08)*, pages 365–380, Las Vegas, Nevada, USA, 15–17 October 2008.

[155] J.F. Sturm. *Using SeDuMi 1.02, A MATLAB toolbox for optimization over symmetric cones, Version 1.05.* Tilburg University, Tilburg, The Netherlands, October 2001.

[156] C.W. Tan, D.P. Palomar, and M. Chiang. Energy-robustness tradeoff in cellular network power control. *IEEE/ACM Transactions on Networking*, 17(3):912–925, June 2009.

[157] P.H. Tan and L.K. Rasmussen. The application of semidefinite programming for detection in CDMA. *IEEE Journal on Selected Areas in Communications*, 19(8):1442–1449, August 2001.

[158] M. Tao, Y.-C. Liang, and F. Zhang. Adaptive resource allocation for delay differentiated traffic in multiuser OFDM systems. In *IEEE International Conference on Communications 2006 (ICC'06)*, volume 10, pages 4403–4408, Istanbul, Turkey, 11-15 June 2006.

[159] Y. T'Joens, T. Bostoen, S. Van den Bosch, and P. Vandaele. Managing home an daccess domains in modern broadband networks. *Bell Labs Technical Journal*, 13(1):247–262, Spring 2008.

[160] P. Tsiaflakis, M. Diehl, and M. Moonen. Distributed spectrum management algorithms for multiuser DSL networks. *IEEE Transactions on Signal Processing*, 56(10):4825–4843, October 2008.

[161] 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*, 58(4):2230–2245, April 2010.

[162] P. Tsiaflakis, C.W. Tan, Y. Yi, M. Chiang, and M. Moonen. Optimality certificate of dynamic spectrum management in multi-carrier interference channels. In *IEEE International Symposium on Information Theory 2008 (ISIT'08)*, pages 1298–1302, Toronto, Canada, 6–11 July 2008.

[163] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden. A low complexity optimal spectrum balancing algorithm for digital subscriber lines. *Signal Processing*, 87(7):1735–1753, July 2007.

[164] P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen. Green DSL: Energy-efficient DSM. In *IEEE International Conference on Communications 2009 (ICC'09)*, Dresden, Germany, 14-18 June 2009.

[165] P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen. Fair greening for DSL broadband access. *ACM SIGMETRICS Performance Evaluation Review*, 37(4):74–78, March 2010.

[166] P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen. Fair greening of broadband access: spectrum management for energy-efficient DSL networks. *to appear in EURASIP Journal on Wireless Communications and Networking*, 2011.

[167] D.E.J. Van Bruyssel. Multi-carrier modem transmitter with controlled transmit signal quality degradation for improving stability of operation. United States Patent No. US 7,991,072 B2, August 2011.

[168] E. Van den Bogaert, T. Bostoen, J. Verlinden, R. Cendrillon, and M. Moonen. ADSL transceivers applying DSM and their nonstationary noise robustness. *EURASIP Journal on Applied Signal Processing*, 2006.

[169] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, March 1996.

[170] F. Vanier. World broadband statistics: Short report Q3 2010. Technical report, Point Topic Ltd, December 2010.

[171] W. Vereecken, W. Van Heddeghem, M. Deruyck, B. Puype, B. Lannoo, W. Joseph, D. Colle, L. Martens, and P. Demeester. Power consumption in telecommunication networks: Overview and reduction strategies. *IEEE Communications Magazine*, 49(6):62–69, June 2011.

[172] J. Verlinden and D. V. Bruyssel. Virtual noise mechanism. in ANSI Contribution NIPP-NAI-2005-049, San Francisco, CA, February 2005.

[173] J. Verlinden, E. Van Den Bogaert, T. Bostoen, R. Cendrillon, and M. Moonen. Protecting the robustness of ADSL and VDSL DMT modems when applying DSM. In *International Zurich Seminar on Communications 2004 (IZS'04)*, Zürich, Switzerland, February 2004.

[174] S.A. Vorobyov, A.B. Gershman, and L. Zhi-Quan. Robust adaptive beamforming using worst-case performance optimization: A solution to the signal mismatch problem. *IEEE Transactions on Signal Processing*, 51(2):313–324, February 2003.

[175] M. Živković, G. Kramer, C.J. Nuzman, C. Posthuma, J. Wheeler, P. Whiting, and A.J. de Lind van Wijngaarden. Performance of digital subscriber line spectrum optimization algorithms. *Bell Labs Technical Journal*, 13(1):129–146, Spring 2008.

[176] X. Wang, G.B. Giannakis, and Y. Yu. Channel-adaptive optimal OFDMA scheduling. In *Annual Conference on Information Sciences and Systems 2007 (CISS '07)*, pages 536–541, Baltimore, MD, USA, 14–16 March 2007.

[177] P.C. Weeraddana, M. Codreanu, M. Latva-aho, and A. Ephremides. Resource allocation for cross-layer utility maximization in wireless networks. *IEEE Transactions on Vehicular Technology*, 60(6):2790–2809, July 2011.

[178] M. Wolkerstorfer, J. Jaldén, and T. Nordström. Low-complexity optimal discrete-rate spectrum balancing in digital subscriber lines. *submitted to Elsevier Signal Processing*, October 2011.

[179] M. Wolkerstorfer, J. Jaldén, and T. Nordström. Column generation for discrete-rate multi-user and multi-carrier power control. *accepted in IEEE Transactions on Communications*, January 2012.

[180] M. Wolkerstorfer and T. Nordström. Coverage optimization in DSL networks by low-complexity discrete spectrum balancing. In *IEEE Global Communications Conference 2011 (Globecom'11)*, Houston, Texas, USA, 5-9 December 2011.

[181] M. Wolkerstorfer and T. Nordström. Heuristics for discrete power control - A case-study in multi-carrier DSL networks. In *ALIO/EURO Workshop on Applied Combinatorial Optimization 2011*, Porto, Portugal, 4–6 May 2011.

[182] M. Wolkerstorfer, T. Nordström, B. Krasniqi, M. Wrulich, and Ch. Mecklenbräuker. OFDM/OFDMA subcarrier allocation. In N. Zorba, Ch. Skianis, and Ch. Verikoukis, editors, *Cross layer designs in WLAN systems*, volume 1, chapter 6, pages 177–216. Troubador Publishing Ltd., 2011.

[183] M. Wolkerstorfer, T. Nordström, and D. Statovci. Delay-constrained scheduling for interference-limited multi-carrier systems. In *IEEE International Workshop on Cross-Layer Design*, Palma de Mallorca, Spain, 11–12 June 2009.

[184] M. Wolkerstorfer, T. Nordström, and D. Statovci. Energy-efficient spectrum balancing in DSL. Technical Report FTW-TR-2010-003, The Telecommunications Research Center Vienna (FTW), Vienna, Austria, May 2010.

[185] M. Wolkerstorfer, B. Putra, T. Nordström, and S. Trautmann. Modeling and optimization of the line-driver power consumption in xDSL systems. *submitted to EURASIP Journal on Advances in Signal Processing*, February 2012.

[186] M. Wolkerstorfer, D. Statovci, and T. Nordström. Dynamic spectrum management for energy-efficient transmission in DSL. In *IEEE International Conference on Communications Systems 2008 (ICCS'08)*, pages 1015–1020, Guangzhou, China, 19–21 November 2008.

[187] M. Wolkerstorfer, D. Statovci, and T. Nordström. Duality-gap bounds for multi-carrier systems and their application to periodic scheduling. In *IEEE International Conference on Communications 2010 (ICC'10)*, Cape Town, South Africa, May 23-27 2010.

[188] M. Wolkerstorfer, D. Statovci, and T. Nordström. Robust spectrum management for DMT-based systems. *IEEE Transactions on Signal Processing*, 58(6):3238 – 3250, June 2010.

[189] M. Wolkerstorfer, D. Statovci, and T. Nordström. Energy-saving by low-power modes in ADSL2. *accepted in Elsevier Computer Networks, Special Issue on Green Communication Networks*, January 2012.

[190] C.Y. Wong, R.S. Cheng, K.B. Letaief, and R.D. Murch. Multiuser OFDM with adaptive subcarrier, bit, and power allocation. *IEEE Journal on Selected Areas in Communications*, 17(10):1747–1758, October 1999.

[191] Wen Xu, C. Schroeder, and P.A. Hoeher. A stochastic MIMO model for far-end crosstalk in VDSL cable binders. In *IEEE International Conference on Communications 2009 (ICC'09)*, Dresden, Germany, 14-18 June 2009.

[192] 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*, 56(7):2875 – 2885, July 2008.

[193] K. Yang, Y. Wu, J. Huang, X. Wang, and S. Verdu. Distributed robust optimization for communication networks. In *IEEE International Conference on Computer Communications 2008 (INFOCOM'08)*, pages 1157–1165, Phoenix, Arizona, USA, 15–17 April 2008.

[194] R.D. Yates. A framework for uplink power control in cellular radio systems. *IEEE Journal on Selected Areas in Communications*, 13(7):1341–1347, September 1995.

[195] G. Young and D. Zimring. ADSL2Plus configuration guidelines for IPTV. Technical Report TR-176 1.00, The Broadband Forum, September 2008.

[196] D.D. Yu, K. Seong, and J.M. Cioffi. Multiuser discrete bit-loading for digital subscriber lines. In *IEEE International Conference on Communications 2007 (ICC'07)*, pages 2755–2760, Glasgow, Scotland, 24–28 June 2007.

[197] W. Yu, G. Ginis, and J. M. Cioffi. Distributed multiuser power control for digital subscriber lines. *IEEE Journal on Selected Areas in Communications*, 20(5):1105–1115, June 2002.

[198] W. Yu and R. Lui. Dual methods for nonconvex spectrum optimization of multicarrier systems. *IEEE Transactions on Communications*, 54(7):1310–1322, July 2006.

[199] Wei Yu, R. Lui, and R. Cendrillon. Dual optimization methods for multiuser orthogonal frequency division multiplex systems. In *IEEE Global Telecommunications Conference 2004 (Globecom'04)*, volume 1, pages 225–229, Dallas, Texas, USA, 29 November – 3 December 2004.

[200] Y. Yu and G.B. Giannakis. Joint congestion control and OFDMA scheduling for hybrid wireline-wireless networks. In *IEEE International Conference on Computer Communications 2007 (INFOCOM'07)*, pages 973–981, Anchorage, Alaska, USA, 6–12 May 2007.

[201] Ch. Zeng, C. Aldana, A.A. Salvekar, and J.M. Cioffi. Crosstalk identification in xDSL systems. *IEEE Journal of Selected Areas in Communications*, 19(8):1488–1496, August 2001.