Column Generation for Discrete-Rate Multi-User and Multi-Carrier Power Control

Martin Wolkerstorfer, *Member, IEEE*, Joakim Jaldén, *Member, IEEE*, and Tomas Nordström, *Senior Member, IEEE*

***Abstract*—We consider a constrained multi-carrier power allocation problem in interference-limited multi-user systems with a finite set of transmission rates. The Lagrange relaxation is a common technique for decomposing such problems into independently solvable per-subcarrier problems. Deviating from this approach our main contribution is the proposal of a novel spectrum management framework based on a Nonlinear Dantzig-Wolfe problem decomposition. It allows for suboptimal initialization and suboptimal power allocation methods with low complexity. While we show that the combinatorial per-subcarrier problems have polynomial complexity in the number of users, we find that such suboptimal methods are indispensable in large systems. Thus we give an overview of various basic dual heuristics and provide simulation results on a set of thousand digital subscriber line (DSL) networks which show the superior performance of our framework compared to previous power control algorithms.**

***Index Terms*—Power control, DSL, optimization methods, interference channels.**

I. INTRODUCTION

IN this paper we study the dynamic spectrum management (DSM) problem of optimal bit and power allocation in interference-limited multi-user and multi-carrier systems with a finite set of transmission rates, focussing on its setting in digital subscriber lines (DSL). Lagrange dual relaxation (LDR), a constrained optimization technique [1], is applicable to the DSM problem for lowering its combinatorial complexity in the number of subcarriers. The application of LDR in power spectrum shaping for improving the spectral compatibility between different DSL technologies has been proposed in [2]. In [3] LDR was applied to a combinatorial multi-user DSM problem in modern DSL systems, which triggered further research on optimal and/or low-complexity LDR-based DSM algorithms for DSL. We refer to [4]–[6] for examples of discrete-rate LDR-based optimization algorithms that are most relevant for our work, and to [7]–[9] for an overview of various continuous optimization schemes for the DSM problem.

An alternative to the LDR is the time-sharing relaxation which allows for a convex combination of various power allocation solutions. Optimal subcarrier and power allocation in orthogonal frequency division multiple access (OFDMA) networks with *continuous* power allocation was recently shown to have a polynomial time approximation [10], where also an equivalent linear time-sharing formulation was derived. Similarly, continuous sharing of subcarriers (in time or frequency) was shown to yield a convex and therefore polynomially solvable optimization problem under continuous rate and power allocation in [11], [12]. The time-sharing problem relaxation is known to be the strong dual problem to the LDR, also when the set of transmission rates is finite [1], [13]. In DSM problems for interference-limited DSL systems 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 [14], [15]. Complexity reduction ideas for the per-subcarrier problems in multi-user systems with a finite set of transmission rates were proposed in [4]–[6], [16], [17].

Our contributions and the outline of the paper are as follows: After introducing the optimization model and its time-sharing relaxation in Section II we will propose a novel framework for multi-carrier power control based on a nonlinear Dantzig-Wolfe (NDW) decomposition and a problem "disaggregation" [18] in Section III. This approach differs from previous Lagrange relaxation schemes for DSM in interference-limited systems [3], [4], [6], [19], [20] 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 III-B 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 and weighted proportional sum-rate fairness. Differently to most work on dual-relaxation based DSM schemes, in Section III-C 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 [19] or a positive duality gap [10]. Furthermore, it bears the potential to use a combination of optimal and low-complexity subopti-

Paper approved by K. K. Leung, the Editor for Wireless Network Access and Performance of the IEEE Communications Society. Manuscript received July 9, 2011; revised November 15, 2011 and February 8, 2012.

M. Wolkerstorfer and T. Nordström are with the FTW Telecommunications Research Center Vienna, Donau-City-Strasse 1, A-1220 Vienna, Austria (e-mail: {wolkerstorfer, nordstrom}@ftw.at). T. Nordström is also with the Centre for Research on Embedded Systems (CERES), Halmstad University, Box 823, 30118 Halmstad, Sweden (e-mail: tomas.nordstrom@hh.se).

J. Jaldén is with the ACCESS Linnaeus Center, Signal Processing Lab, KTH Royal Institute of Technology, Osquldas väg 10, 100 44, Stockholm, Sweden (e-mail: jalden@kth.se).

This work was supported by grant ICA08-0046 of the Swedish Foundation for Strategic Research (SSF). The Competence Center FTW Forschungszentrum Telekommunikation Wien GmbH is funded within the program COMET - Competence Centers for Excellent Technologies by BMVIT, BMWA, and the City of Vienna. The COMET program is managed by the FFG.

Digital Object Identifier 10.1109/TCOMM.2012.070912.110444

mal 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, an exception being for instance the scheme in [21] which uses approximate subgradients [1].

The applicability to various DSM problems and the possibility to solve the subproblems heuristically and in parallel are the most attractive features of our framework. This is because we found the optimal subcarrier solution for general DSL networks intractable to compute [22], although we will show that these subproblems do in fact have polynomial complexity in the number of users. Hence, we devote Section IV to a brief overview of two basic building blocks for more sophisticated meta-heuristics [23]: the greedy search and the local search. As an example, for scenarios with few dominant interfering users we propose a greedy search heuristic which can explicitly make use of this network feature. The simulation results following in Section V demonstrate the advantage of performing a heuristic combinatorial search jointly for all users. Besides a 50 user very high speed DSL (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. Our contributions are summarized in Section VI.

II. Model and Problem Formulations

We assume a far-end crosstalk limited DSL system with U coordinated lines employing discrete multi-tone (DMT) modulation and denote the sets of indices for users and subcarriers by $\mathcal{U} = \{1, \ldots, U\}$ and $\mathcal{C} = \{1, \ldots, C\}$, respectively. Under standard assumptions [24] the achievable rate per DMT-symbol for user $u \in \mathcal{U}$ on subcarrier $c \in \mathcal{C}$ can be approximated by

$$r_c^u(\mathbf{p}_c) = \log_2\Big(1 + \frac{H_c^{uu} p_c^u}{\Gamma(\sum_{i \in \mathcal{U} \setminus u} H_c^{ui} p_c^i + N_c^u)}\Big), \quad (1)$$

where $\mathbf{p}_c = [p_c^1, \ldots, p_c^U]^T$, p_c^u is the transmit power spectral density (PSD), where Γ is the SNR-gap to capacity, where N_c^u is the total receiver noise spectral density, and where H_c^{uu} and H_c^{ui} 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 $\mathbf{r}_c(\mathbf{p}_c) = [r_c^1(\mathbf{p}_c), \ldots, r_c^U(\mathbf{p}_c)]^T$ and use $\mathbf{p}_c(\mathbf{r}_c)$ to denote the unique [25] power allocation resulting in the rate vector $\mathbf{r}_c$.

A. Original Optimization Problem

A generic multi-user DSM problem in the form of a multi-dimensional nonlinear Knapsack problem [26] can be formulated as

$$P^*_{(\mathbf{R},\mathbf{P})} = \underset{\mathbf{p}_c \in \mathcal{Q}_c, c \in \mathcal{C}}{\text{minimize}} \quad \sum_{c \in \mathcal{C}} f_c(\mathbf{p}_c, \hat{\mathbf{w}}, \check{\mathbf{w}}) \quad (2a)$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} \mathbf{r}_c(\mathbf{p}_c) \succeq \mathbf{R}, \quad (2b)$$

$$\sum_{c \in \mathcal{C}} \mathbf{p}_c \preceq \mathbf{P}, \quad (2c)$$

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

$$f_c(\mathbf{p}_c, \hat{\mathbf{w}}, \check{\mathbf{w}}) = \hat{\mathbf{w}}^T \mathbf{p}_c - \check{\mathbf{w}}^T \mathbf{r}_c(\mathbf{p}_c), \ c \in \mathcal{C}, \quad (3)$$

where $\hat{\mathbf{w}}, \check{\mathbf{w}} \in \mathcal{R}_+^U$ are weights, where $\mathbf{R} \in \mathcal{R}_+^U$ are the target-rates in [bits/DMT-symbol], where $\mathbf{P} \in \mathcal{R}_+^U$ are the maximum sum-powers, and where

$$\mathcal{Q}_c = \{\mathbf{p}_c | r_c^u(\mathbf{p}_c) \in \mathcal{B}, \ 0 \le p_c^u \le \hat{p}_c^u, \forall u \in \mathcal{U}\}, \quad (4)$$

is the set of feasible PSD's on subcarrier c. In (4), $\hat{p}_c^u$ are spectral mask constraints and $\mathcal{B} = \{0, \Delta, 2\Delta, \ldots, \hat{B}\}$ is the set of positive, discrete bit allocations per-subcarrier, where we assume rate-steps of equal size Δ and a single maximum number of loaded bits $\hat{B}$ for all users, respectively. 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. For later reference we define the dual problem to (2) as

$$D^*_{(\mathbf{R},\mathbf{P})} = \underset{\boldsymbol{\lambda} \succeq \mathbf{0}, \boldsymbol{\nu} \succeq \mathbf{0}}{\text{maximize}} \quad q^{\text{tot}}(\boldsymbol{\lambda}, \boldsymbol{\nu}), \quad (5)$$

where the dual function [1] is defined as

$$q^{\text{tot}}(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \sum_{c \in \mathcal{C}} q_c(\boldsymbol{\lambda}, \boldsymbol{\nu}) + \boldsymbol{\lambda}^T \mathbf{R} - \boldsymbol{\nu}^T \mathbf{P}, \quad (6)$$

where $\boldsymbol{\lambda}, \boldsymbol{\nu} \in \mathcal{R}^U$ are the Lagrange multipliers associated with constraints (2b) and (2c), respectively, and where

$$q_c(\boldsymbol{\lambda}, \boldsymbol{\nu}) = \min_{\mathbf{p}_c \in \mathcal{Q}_c} \{f_c(\mathbf{p}_c, \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda})\}, \quad \forall c \in \mathcal{C}. \quad (7)$$

B. Time-Sharing Relaxation

As an alternative relaxation to the dual relaxation in (5) a problem formulation involving time-sharing among various bit and power allocations was suggested for OFDMA systems in [11], [12] and for interference-limited systems in [14], [15]. For time-sharing[1] 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 $0 \le \xi_c^i \le 1$ 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 (2) as the linear program (LP)

$$P^{*,\text{ts}}_{(\mathbf{R},\mathbf{P})} = \underset{\xi_c^i \ge 0, i \in \mathcal{I}_c, c \in \mathcal{C}}{\text{minimize}} \quad \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}_c} f_c(\mathbf{p}_c^i, \hat{\mathbf{w}}, \check{\mathbf{w}}) \xi_c^i \quad (8a)$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}_c} \mathbf{r}_c(\mathbf{p}_c^i) \xi_c^i \succeq \mathbf{R}, \quad (8b)$$

$$\sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}_c} \mathbf{p}_c^i \xi_c^i \preceq \mathbf{P}, \quad (8c)$$

$$\sum_{i \in \mathcal{I}_c} \xi_c^i = 1, \quad \forall c \in \mathcal{C}, \quad (8d)$$

where constraint (8d) ensures that the time-shares on each subcarrier sum to one. We emphasize that the problem in (8) allows for inter-user interference using a finite set of rates according to (1), but varying the fraction of time the possible power combinations on each subcarrier are applied. While we

[1] Note that we regard time-sharing solely as an algorithmic detour to obtain solutions for our original problem in (2), as will become clearer in Section III.

employed time-sharing on a per-subcarrier basis, note that it can also be performed in "aggregated" form [18] 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 (8):

Theorem 1: Define $\mathcal{I}_c^{\boldsymbol{\xi}} = \{i \in \mathcal{I}_c | \xi_c^i > 0\}$, for $c \in \mathcal{C}$, and the set of subcarriers where time-sharing occurs as $\mathcal{C}_{\boldsymbol{\xi}}^+ = \{c \in \mathcal{C} | \, |\mathcal{I}_c^{\boldsymbol{\xi}}| \geq 2\}$. Assuming feasibility of (5) it holds that $P_{(\mathbf{R},\mathbf{P})}^{*,\text{ts}} = D_{(\mathbf{R},\mathbf{P})}^{*}$ and there exists a solution $\tilde{\boldsymbol{\xi}}$ to (8) with $|\mathcal{C}_{\tilde{\boldsymbol{\xi}}}^+| \leq 2U$, i.e., time-sharing is required on at most $2U$ subcarriers.

Proof: The proof of the first statement is analogous to that explicitly given in [13] and follows from showing the equivalence between the dual linear program to (8) and the dual problem in (5), and strong duality due to the assumed feasibility of (5). The second statement follows as (8) has $2U + C$ constraints and therefore a solution exists[2] that has at most this number of non-zero variables [27]. As $|\mathcal{I}_c^{\boldsymbol{\xi}}| \geq 1$, 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 [28, p. 374], cf. the geometric interpretations of (2) and (5) in [13]. ■

For completeness, we proceed by analyzing the scalability of the time-sharing problem in (8). The number of feasible bit and power allocations $|\mathcal{Q}_c|, c \in \mathcal{C}$, and therefore the number of variables in (8) grows with an increasing number of users U. However, the following result indicates how interference among users restricts this growth.

Theorem 2: Assuming $\frac{H_c^{ui}}{H_c^{uu}} \geq \alpha > 0$, the number of feasible allocations $|\mathcal{Q}_c|$ for subcarrier $c \in \mathcal{C}$ grows at most polynomially as $O(U^{\hat{U}})$, where the constant exponent $\hat{U}$ is given by

$$\hat{U} = 1 + \left(\Gamma(2^{\Delta} - 1)\alpha\right)^{-1}. \tag{9}$$

See Appendix A for a proof.

The parameter α has the interpretation of a minimal normalized cross-channel attenuation coefficient in the network.[3] Using this minimal value we obtain a lower bound for the (normalized) interference noise per interfering user in (1). Assuming all users transmit at a positive rate we thereby obtain an upper-bound for the number of users that can be supported by the system in (9), hence limiting the growth of the time-sharing formulation in (8) as shown by Theorem 2.

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., $\alpha < -11.3\,\text{dB}$ on typical VDSL lines [29]. In Fig. 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 $l_u = 800\,\text{m}, \forall u \in \mathcal{U}$, and simulation parameters as specified in Section V. Additionally we show a polynomial of degree $\hat{U}$ shifted by $\hat{U}$, cf. (19c). 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 (8) grows with the number of users U, this number (although of complexity theoretic interest) is too large for a direct LP solution. Hence, we proceed in the next section by analyzing a decomposition scheme which works with a small subset of these time-sharing variables.

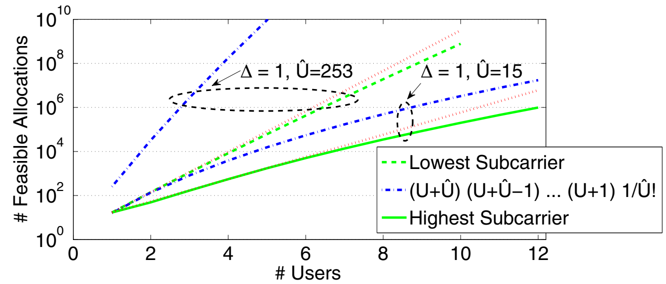

Fig. 1. Number of feasible bit allocations $|\mathcal{Q}_c|$ on selected subcarriers $c \in \mathcal{C}$ in a symmetric VDSL scenario with 800 m long lines.

III. A Novel Framework for DSM

In this section we propose a novel DSM framework which approaches the original problem in (2) by iteratively optimizing the time-sharing relaxation in (8), 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 (7) as shown in Section III-A, its applicability to various DSM objectives as highlighted in Section III-B, and the possibility for heuristic solutions of the per-subcarrier problems as studied in Section IV. Our method is related to previous dual relaxation based DSM algorithms [3], [4], [6], [19], [20] through Theorem 1 and partly motivated by the results in [4], [10] showing a vanishing duality gap — that is the difference between the optimal objectives in the original problem in (2) and its dual in (5).

A. Nonlinear Dantzig-Wolfe Decomposition

The decomposition scheme described next from first principles is based on the mathematical programming concept of column[4] generation [30], or more precisely a nonlinear Dantzig-Wolfe (NDW) decomposition [18], [31, Ch. 23] of (2). Applications of this decomposition approach in the area of wireless communication can be found in [32]–[35]. At iteration t of the algorithm we consider a subset of all columns

[2]Such a solution is referred to as a "basic solution" in LP theory [27].

[3]We acknowledge the fact that theoretically there are scenarios where such a lower bound can not be assumed.

[4]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 (8).

$\mathcal{I}_c^{(t)} \subseteq \mathcal{I}_c, c \in \mathcal{C}$ in (8), yielding the restricted master problem

$$P^{*,ts(t)}_{(\mathbf{R},\mathbf{P})} = \tag{10a}$$

$$\underset{\xi' \geq 0, \xi_c^i \geq 0, i \in \mathcal{I}_c^{(t)}, c \in \mathcal{C}}{\text{minimize}} \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}_c^{(t)}} f_c(\mathbf{p}_c^i, \hat{\mathbf{w}}, \check{\mathbf{w}}) \xi_c^i + f'(\mathbf{P}, \mathbf{R}) \xi'$$

$$\text{subject to} \quad \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}_c^{(t)}} \mathbf{r}_c(\mathbf{p}_c^i) \xi_c^i + \mathbf{R}\xi' \succeq \mathbf{R}, \tag{10b}$$

$$\sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}_c^{(t)}} \mathbf{p}_c^i \xi_c^i + \mathbf{P}\xi' \preceq \mathbf{P}, \tag{10c}$$

$$\sum_{i \in \mathcal{I}_c^{(t)}} \xi_c^i + \xi' = 1, \quad \forall c \in \mathcal{C}, \tag{10d}$$

where we added an artificial column with weight ξ' and cost $f'(\mathbf{P}, \mathbf{R}) = \hat{\mathbf{w}}^T(\mathbf{P} + \boldsymbol{\delta}) - \check{\mathbf{w}}^T(\mathbf{R} - \boldsymbol{\delta})$, for some arbitrary $\boldsymbol{\delta} \succ \mathbf{0}$. By setting $\xi' = 1$ and $\xi_c^i = 0$, for all $i \in \mathcal{I}_c^{(t)}$ and $c \in \mathcal{C}$, it can be seen that this "aggregated" column makes (10) always feasible. Furthermore, the choice of cost leads to the following results, indicating that this artificial column does not alter the solution when (8) is feasible.

Theorem 3: At the optimum of (10) we have $\xi' = 0$ if a feasible solution $\{\xi_c^i\}$ to (8) exists with $\xi_c^i = 0$, for all $i \in \mathcal{I}_c \setminus \mathcal{I}_c^{(t)}, c \in \mathcal{C}$, and $\xi' = 1, \xi_c^i = 0$, for all $i \in \mathcal{I}_c^{(t)}, c \in \mathcal{C}$, otherwise.

See Appendix B for a proof.

Corollary 1: Assuming feasibility of (8) we have $P^{*,\text{ts}(t)}_{(\mathbf{R},\mathbf{P})} \geq P^{*,\text{ts}}_{(\mathbf{R},\mathbf{P})}$.

Proof: By Theorem 3 we either have $\xi' = 0$ or $\xi' = 1$ at the optimum of (10). The corollary follows as $\mathcal{I}_c^{(t)} \subseteq \mathcal{I}_c, \forall c \in \mathcal{C}$, and by $\boldsymbol{\delta} \succ \mathbf{0}$ and feasibility in (8) any solution of (8) has a lower objective than $f'(\mathbf{P}, \mathbf{R})$. ■

After solving (10), the second task at each iteration in a column generation scheme is to compute new columns to be added to the master problem in (10) in order to reduce the gap described in Corollary 1. Relaxing constraints (10b), (10c) and (10d) in the restricted master problem at iteration t and denoting their Lagrange multipliers by $\boldsymbol{\lambda}^{(t)}, \boldsymbol{\nu}^{(t)} \in \mathcal{R}_+^U$ and $\boldsymbol{\Phi}^{(t)} \in \mathcal{R}^C$, respectively, and also including variables $\xi_c^i = 0$, for all $i \in \mathcal{I}_c \setminus \mathcal{I}_c^{(t)}$ and $c \in \mathcal{C}$, we can write the Lagrangian for (10) as

$$L^{(t)} = \sum_{c \in \mathcal{C}} \sum_{i \in \mathcal{I}_c} \left(f_c(\mathbf{p}_c^i, \hat{\mathbf{w}} + \boldsymbol{\nu}^{(t)}, \check{\mathbf{w}} + \boldsymbol{\lambda}^{(t)}) + \Phi_c^{(t)} \right) \xi_c^i + \left(\mathbf{R}^T \boldsymbol{\lambda}^{(t)} - \mathbf{P}^T \boldsymbol{\nu}^{(t)} - \sum_{c \in \mathcal{C}} \Phi_c^{(t)} \right)(1 - \tilde{\xi}) + \tilde{f}(\mathbf{P}, \mathbf{R}) \tilde{\xi}. \tag{11}$$

Adding any column $i \in \mathcal{I}_c \setminus \mathcal{I}_c^{(t)}$ to (10) with negative derivative[5] $\partial L^{(t)} / \partial \xi_c^i = f_c(\mathbf{p}_c^i, \hat{\mathbf{w}} + \boldsymbol{\nu}^{(t)}, \check{\mathbf{w}} + \boldsymbol{\lambda}^{(t)}) + \Phi_c^{(t)}$ at the (dual) optimum of (10) lowers the optimal objective of (10) or leaves it unchanged.[6] Hence, a simple criterion is to pick the column on subcarrier c with minimal derivative, leading to decomposable subproblems similar to (7) in the form of

$$q_c^{\text{red}}(\boldsymbol{\lambda}^{(t)}, \boldsymbol{\nu}^{(t)}, \Phi_c^{(t)}) = q_c(\boldsymbol{\lambda}^{(t)}, \boldsymbol{\nu}^{(t)}) + \Phi_c^{(t)}, \tag{12a}$$

$$= \min_{i \in \mathcal{I}_c} \left\{ f_c(\mathbf{p}_c^i, \hat{\mathbf{w}} + \boldsymbol{\nu}^{(t)}, \check{\mathbf{w}} + \boldsymbol{\lambda}^{(t)}) \right\} + \Phi_c^{(t)}, \forall c \in \mathcal{C}. \tag{12b}$$

Corollary 2 (of Thm. 2): The per-subcarrier subproblem in (12b) has polynomial complexity in the number of users U.

Proof: By definition $|\mathcal{I}_c| = |\mathcal{Q}_c|$, where $|\mathcal{Q}_c|$ has polynomial size in U by Theorem 2, and the evaluation of the objective in (3) has polynomial complexity. It remains to be shown that there exists an algorithm for enumerating $\mathcal{Q}_c$ with polynomial complexity in U. To do so we interpret bit-loading as a search in a search-tree of depth U [36]. A depth-first search begins with an allocation $\mathbf{0}$ and sequentially proceeds to higher bits, starting at level U. Everytime an infeasible allocation is encountered it returns to the next-lower tree level. Therefore, while testing all feasible allocations $\mathbf{p}_c \in \mathcal{Q}_c$ (and therefore being optimal) this search never tests more than $U|\mathcal{Q}_c|$ infeasible allocations $\mathbf{p}_c \notin \mathcal{Q}_c$, which concludes the proof. ■

The proposed DSM algorithm iterates between solving (10) and the C subproblems in (12b). Hence, the number of columns in (10) increases by at most C in each iteration, cf. Algorithm 1. We emphasize that any potentially suboptimal solution to (12b) with negative objective may improve the restricted master problem (10), which is the reason why problem (12b) is amenable for fast heuristics, cf. Line 7 in Algorithm 1 and our overview on basic heuristics in Section IV. Furthermore, we have that if $q_c(\boldsymbol{\lambda}^{(t)}, \boldsymbol{\nu}^{(t)}) \geq -\Phi_c^{(t)}, \forall c \in \mathcal{C}$, then $P^{*,ts(t)}_{(\mathbf{R},\mathbf{P})} = P^{*,\text{ts}}_{(\mathbf{R},\mathbf{P})}$, cf. (12a). The same conclusion can be drawn if, $\forall c \in \mathcal{C}$, we have $i \in \mathcal{I}_c^{(t)}$, where i is the minimum argument in (12b) [13], [33]. This means that the algorithm terminates if not at least one new allocation on any subcarrier is added to the master problem in (10). A finite convergence time of the algorithm then follows from the finiteness of $|\mathcal{Q}_c|, c \in \mathcal{C}$. The negativity of $q_c^{\text{red}}(\boldsymbol{\lambda}^{(t)}, \boldsymbol{\nu}^{(t)}, \Phi_c^{(t)})$ as a necessary criterion for an improving column can also be exploited on each subcarrier to reduce the complexity of solving (12b), for instance by using $-\Phi_c^{(t)}$ as the initial incumbent objective used for pruning the search tree in branch-and-bound based algorithms, cf. [36] for an application.

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

$$\tilde{\zeta} = P^{*,\text{ts}(t)}_{(\mathbf{R},\mathbf{P})} - \max_{1 \leq j \leq t} \{ q^{\text{tot}}(\boldsymbol{\lambda}^{(j)}, \boldsymbol{\nu}^{(j)}) \}, \tag{13}$$

where $q^{\text{tot}}(\cdot, \cdot)$ is defined in (6), 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 (10) to reduce its size.

B. Further Properties and Applications of NDW-DSM

We claim that the presented disaggregated NDW-DSM algorithm is numerically more stable than previous Lagrange

[5]This derivative is also referred to as the "reduced cost" of a column [27].

[6]More precisely, assuming non-degeneracy of a basic solution of (10) one can pivot on the new variable ξ_c^i with $\partial L^{(t)} / \partial \xi_c^i < 0$ and thereby maintain feasibility while *strictly* decreasing the objective value [27]. From a dual perspective, non-degeneracy corresponds to uniqueness of the dual solution $\boldsymbol{\lambda}^{(t)}, \boldsymbol{\nu}^{(t)}, \Phi_c^{(t)}, c \in \mathcal{C}$, to (10) [27]. Under this uniqueness it follows from the existence of a negative gradient direction w.r.t. (11) at $\boldsymbol{\lambda}^{(t)}, \boldsymbol{\nu}^{(t)}, \Phi_c^{(t)}, c \in \mathcal{C}$, and strict duality that the optimal objective of (10) decreases *strictly*.

Algorithm 1 NDW-DSM

1: Initialize $t = 1$, $\mathcal{I}_c^{(1)}, \forall c \in \mathcal{C}, \tilde{\zeta} = \infty$, $P_{(\mathbf{R},\mathbf{P})}^{*,ts(0)} = \infty$, $\bar{\delta}$
2: **while** $\tilde{\zeta} > \zeta^{\text{tgt}}$ and $|P_{(\mathbf{R},\mathbf{P})}^{*,ts(t-1)} - P_{(\mathbf{R},\mathbf{P})}^{*,ts(t)}| > \bar{\delta}\,|P_{(\mathbf{R},\mathbf{P})}^{*,ts(t-1)}|$ **do**
3: Solve (10) by a primal-dual LP solver, obtaining dual multipliers $[\boldsymbol{\lambda}^{(t)}, \boldsymbol{\nu}^{(t)}, \boldsymbol{\Phi}^{(t)}]$
4: Obtain new allocations solving (12b) either ...
5: ... a) Optimally (e.g., by branch-and-bound [6], [22])
6: Compute $\tilde{\zeta}$ as in (13)
7: ... b) Suboptimally (initially, e.g., using Algorithm 4)
8: Add only new allocations $i \in \mathcal{I}_c \backslash \mathcal{I}_c^{(t)}$ to (10), $t = t+1$
9: Apply Algorithm 2 to recover solutions to (2)

Algorithm 2 Combination Heuristic for Time-sharing (CHET)

1: $[\{i_c^*\}_{c\in\mathcal{C}}] = \text{CHET}(\{\mathcal{I}_c^{(t)}, \xi_c^i, \forall i \in \mathcal{I}_c^{(t)}\}_{c\in\mathcal{C}}, \boldsymbol{\lambda}, \boldsymbol{\nu})$
2: Initialize $\delta, \kappa \geq 0$, $i_c^* = \text{argmax}_{i\in\mathcal{I}_c^{(t)}}\{\xi_c^i\}, \forall c \in \mathcal{C}$
3: **while** No feasible solution to (2) 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** $\mathbf{P}^* = \sum_{c\in\mathcal{C}} \mathbf{p}^{i_c^*}$, $\mathbf{R}^* = \sum_{c\in\mathcal{C}} \mathbf{r}_c(\mathbf{p}^{i_c^*})$
6: $\breve{\mathcal{I}}_c = \{i \in \mathcal{I}_c^{(t)} \mid f_c(\mathbf{p}_c^i, \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda}) \leq f_c(\mathbf{p}_c^{i_c^*}, \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda}) + (|f_c(\mathbf{p}_c^{i_c^*}, \hat{\mathbf{w}} + \boldsymbol{\nu}, \check{\mathbf{w}} + \boldsymbol{\lambda})| + \kappa) \cdot \delta\}$
7: $\Delta R_u^i = [R_u - R_u^* + r_c^u(\mathbf{p}^{i_c^*}) - r_c^u(\mathbf{p}^i)]_+, \forall u \in \mathcal{U}, i \in \breve{\mathcal{I}}_c$
8: $\Delta P_u^i = [P_u^* - p_u^{i_c^*} + p_u^i - P_u]_+, \forall u \in \mathcal{U}, i \in \breve{\mathcal{I}}_c$
9: $i_c^* = \text{argmin}_{i\in\breve{\mathcal{I}}_c}\{(\hat{\mathbf{w}} + \boldsymbol{\nu})^T \Delta \boldsymbol{P}^i + (\check{\mathbf{w}} + \boldsymbol{\lambda})^T \Delta \boldsymbol{R}^i\}$

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 [10], [19] which may arise due to non-convexity of the original problem in (2). In general, computing a primal feasible allocation based on the dual optimum is then again NP-hard, but there exists a time-sharing solution having an objective value equal to the optimal dual one [10], [13]. This problem was also tackled in [19], [20] by a specific Lagrange multiplier search scheme which similarly yields time-shared solutions but works differently with an aggregated formulation and necessitates optimal per-subcarrier allocation schemes. We point out that the presented basic NDW-DSM scheme may be further improved by stabilization techniques but deem their treatment beyond the scope of this text, cf. [30] 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 (12b). Furthermore, the NDW master problem in (10) can be initialized with the solution obtained from suboptimal algorithms such as iterative spectrum balancing [4], [5] assuming the per-subcarrier feasibility in (4) 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 (10). The memory of per-subcarrier solutions in the NDW master problem in (10) may also pay off in cases where the original problem in (2) needs to be solved for various values of target-rates $\mathbf{R}$ or sum-powers $\mathbf{P}$. As exemplified in [22] sufficiently good solutions may be obtained by first generating a set of columns $\mathcal{I}_c^{(t)}, c \in \mathcal{C}$, for initial target-rates and sum-power values and then repeatedly solving the LP in (10) with this column set but under different target-rate and power constraints. Hence, one solves several LPs instead of repeatedly solving the original problem in (2).

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 (3). 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 (12b). Examples include the maximization of the users' minimum rate or weighted proportional sum-rate fairness. To see this consider the latter objective given by [37] $\sum_{u\in\mathcal{U}} w_u \log(t_u)$, 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 (2b). It can readily be verified following the same steps as in Section III-A that NDW decomposition results in a convex master problem over the same variables as in (10) and additionally the U auxiliary variables $\mathbf{t}$, as well as decomposable subproblems identical in form to those in (12b). We refer to [38] for further optimization problems to which NDW-DSM is applicable.

C. *Combination Heuristic for Time-Sharing (CHET)*

Algorithm 1 targets the solution of the time-shared problem (8), having the same optimal objective as the dual problem (5). In most of the previous work on dual-relaxation based DSM algorithms the problem of recovering feasible solutions for the original problem (2) is either overlooked or circumvented by proposing direct implementations of time-sharing solutions [4], [10], [15], [19], with an exception being [19], [20]. However, we found that the heuristics in [19], [20], 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 V-A for an example. More precisely, the scheme in [20] 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 $\mathcal{I}_c^{(t)} \subseteq \mathcal{I}_c$ are available, enumerating the product set $\prod_{c\in\mathcal{C}} \mathcal{I}_c^{(t)}$ in order to find feasible allocations for (2) remains intractable in general. The suggested algorithm hence iteratively and greedily selects a single allocation $i_c^* \in \mathcal{I}_c^{(t)}, \forall c \in \mathcal{C}$. The main target is feasibility in (2), which is why in Lines 7 and 8 of Algorithm 2 the impact of choosing an allocation $i \in \mathcal{I}_c^{(t)}$ on the sum-power and the sum-rate constraints in

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** $\exists \mathbf{p}\in\mathcal{Q}|r_u(\mathbf{p})=r_u+\Delta, r_i(\mathbf{p})=r_i, \forall i\in\mathcal{U}\setminus\{u\}$,
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: $u^*=\text{argmin}_{u=U^{\text{opt}}+1\ldots U}\ \delta_u,\ \delta^*=\delta_{u^*}$
9: **if** $\delta^*\leq 0$ **then** $r_{u^*}=r_{u^*}+\Delta,\ f^{\text{prev}}=f^{\text{prev}}+\delta_{u^*}$

Algorithm 4 Sequential Greedy Optimization (SEGO)

1: $[\mathbf{r},\mathbf{p}, f(\mathbf{p},\hat{\mathbf{w}}+\boldsymbol{\nu},\check{\mathbf{w}}+\boldsymbol{\lambda})] = \text{SEGO}(\mathbf{r}^0, U^{\text{opt}},\boldsymbol{\lambda},\boldsymbol{\nu})$
2: $\mathbf{r}=\mathbf{r}^0,\ U^{\text{greed}}=U-U^{\text{opt}}$
3: Determine sequence $\mathbf{s}\in\mathcal{R}^{U^{\text{greed}}}$ by ordering users $U^{\text{opt}}+1\leq u\leq U$ in descending order of $(\check{w}_u+\lambda_u)/(\hat{w}_u+\nu_u)$
4: **for** $u=s_1,\ldots,s_{U^{\text{greed}}}$ **do**
5: $[r^u,\mathbf{p}]=\underset{\{r_u\in\mathcal{B},\mathbf{p}\in\mathcal{Q}|\mathbf{r}(\mathbf{p})\succeq\mathbf{r}\}}{\text{argmin}}\{f(\mathbf{p},\hat{\mathbf{w}}+\boldsymbol{\nu},\check{\mathbf{w}}+\boldsymbol{\lambda})\}$

(2b) and (2c) 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 $P^{*,\text{ts}}_{(\mathbf{R},\mathbf{P})}$ by restricting this selection to a subset of columns $\breve{\mathcal{I}}_c\subseteq\mathcal{I}_c^{(t)}, c\in\mathcal{C}$, with objective values in (7) below a certain threshold. This restriction is successively relaxed in case no allocation feasible in (2) was found.

In the following we complete this work by giving an overview of various heuristics for the discrete per-subcarrier problems in (12b) as used in Line 7 of 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 the paper for ease of notation.

IV. Heuristics for Discrete Power Control

In [36] we study the optimal solution of the subproblems in (12b) by branch-and-bound type of algorithms under problem-specific variable-range reduction strategies. Under these strategies we observe that problem instances with as many as 16 users can be solved optimally in short time when crosstalk levels 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: the greedy search and the local search. We refer to [22], [23] for extensions and more details on heuristics for discrete-rate power control.

A. Greedy Search Schemes

In Algorithm 3 we describe an iterative joint greedy optimization scheme (JOGO) for optimizing the per-subcarrier optimization problems (12b) 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 Δ bits for any of the U^{greed} users is calculated and Δ bits are allocated to the user minimizing this cost, cf. Lines 4-9 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 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 $(\hat{w}_u+\nu_u)/(\check{w}_u+\lambda_u)$, 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 $\mathbf{p}(\mathbf{r})$ [16] we have that JOGO has a complexity per loaded bit-step Δ of $O(U^3)$, while that of SEGO is only $O(U^2)$, cf. the cost-update in Lines 4 to 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 [6], [36], 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 $\mathcal{Q}$. 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 $r_u, 1\leq u\leq U^{\text{opt}}$, a heuristic algorithm is used to allocate bits to the remaining $U^{\text{greed}}=U-U^{\text{opt}}$ users. 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 [36] and SEGO in Algorithm 4, respectively. In Section V-A 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.

B. Local Search

An essential part of many meta-heuristics is the local search, where in [23] 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 $\mathbf{r}^{(t)}$ in iteration t of the search to an improving allocation $\mathbf{r}^{(t+1)}\in\mathcal{N}(\mathbf{r}^{(t)})\subseteq\mathcal{B}$ where $f(\mathbf{r}^{(t+1)})<f(\mathbf{r}^{(t)})$. 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. The set $\mathcal{N}(\mathbf{r})$ is called the neighborhood of $\mathbf{r}$, where a simple but effective choice was found to be the set [23]

$$\mathcal{N}^{(2)}(\mathbf{r}) = \mathcal{N}^{(1)}(\mathbf{r})\cup\bar{\mathcal{N}}^{(2)}(\mathbf{r}), \tag{14a}$$

$$\mathcal{N}^{(1)}(\mathbf{r}) = \{\tilde{\mathbf{r}}\in\prod_{u\in\mathcal{U}}\mathcal{B}\mid \tilde{r}_u=r_u\pm\Delta, \tilde{r}_i=r_i,\forall i\in\mathcal{U}\setminus\{u\}, u\in\mathcal{U}\}, \tag{14b}$$

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

$$\tilde{r}_i = r_i, \forall i \in \mathcal{U} \setminus \{u, \bar{u}\}, u \neq \bar{u}, u, \bar{u} \in \mathcal{U}\}, \quad (14c)$$

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 led to a "warm-start" local search scheme in [22]. The following result characterizes the asymptotic size of the proposed neighborhood in U and the complexity of local search, respectively.

Theorem 4: 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 (14b), while the second part on $|\mathcal{N}^{(2)}(\mathbf{r})|$ follows from the proof of Theorem 2 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 $\mathbf{p}(\mathbf{r}) \in \mathcal{Q}$ was shown to be 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 (14c) comprising all allocations generated by changing *exactly* two elements of $\mathbf{r}$ is bounded by $\hat{U}U$. Altogether the size $|\mathcal{N}^{(2)}(\mathbf{r})| = |\mathcal{N}^{(1)}(\mathbf{r})| + |\bar{\mathcal{N}}^{(2)}(\mathbf{r})|$ of the set in (14a) grows linearly in U. ■

Corollary 3 (of Thm. 2 and Thm. 4): The local search for the problem in (12b) has polynomial complexity in U under the assumptions and the neighborhood set of Theorem 4.

Proof: This follows from Theorems 2 and 4 and by the polynomial complexity of evaluating $\mathbf{p}(\mathbf{r})$ by solving a linear system, cf. [16]. ■

The assumptions in Theorem 4 are satisfied in the following. More generally they hold in all randomized schemes in [22], [23] as these only evaluate the neighborhood around rates $\mathbf{r}$ where the power allocation $\mathbf{p}(\mathbf{r})$ is in the feasible set $\mathcal{Q}$.

V. Simulation Results

In Section V-A 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 V-B 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 V-C. Throughout we assume a DSL system with simulation parameters chosen in accordance with the ETSI VDSL standard [39], with $\Gamma = 12.8\,\text{dB}$, $\hat{B} = 15$, $\Delta = 1$, two transmission bands as defined in band plan 997-M1x-M, and with noise comprising alien crosstalk according to ETSI VDSL noise A added to a flat background noise at $-140\,\text{dBm/Hz}$. The channel model is based on the "TP100" cable model in [39], and the common 99 % worst-case crosstalk channel model for European cables [24], cf. [40] for a publicly available DSL simulator. We will compare our algorithm to single-user bit-loading [41] under worst-case crosstalk noise $\sum_{i \in \mathcal{U} \setminus u} H^{ui}_c \hat{p}^i_c, c \in \mathcal{C}$, for user $u \in \mathcal{U}$ ("Mask-Based"), to an iterative spectrum balancing (ISB) algorithm[7], and the central discrete bit-loading (CDBL) algorithm in [16].

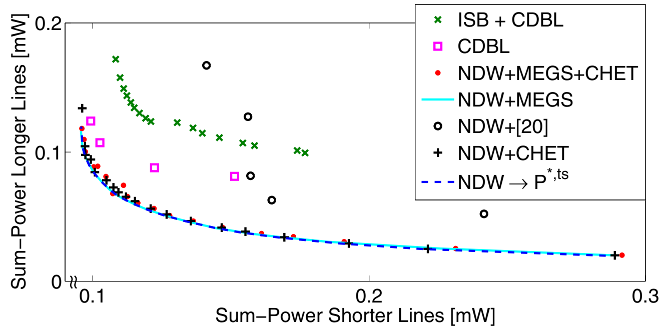

Fig. 2. Sum-power regions in a 4 user VDSL upstream near-far scenario.

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

In Fig. 2 we present the power-regions of various DSM algorithms in a 4 user VDSL upstream near-far scenario.[8] In this scenario we have that the crosstalk between the longer two lines is low compared to the shorter two lines, and we therefore expect the MEGS heuristic presented in Section IV-A to be useful. As Fig. 2 shows, in this near-far scenario the power-region obtained by the MEGS heuristic is close to the lower-bound given by the optimum[9] $P^{*,\text{ts}}$ of the time-sharing problem in (8), 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 CDBL is more than 30 % above that obtained by the combination of NDW-DSM with MEGS and CHET, while applying CDBL 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.[10]

The share of subcarriers where time-sharing is applied at the solutions of the time-sharing problem in (8) is in the range between 0 % (meaning that a single solution was found and the relaxation gap between the problems in (2) and (8) is hence zero) and 60 %. These numbers depend however on the used

[7]In ISB we perform the Lagrange multiplier update sequentially over users while fixing PSDs as in [5] 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 CDBL from this initialization, cf. [42] 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 $30 \leq t \leq 60$ and set $\bar{\delta}$ to 0.1 ppm, cf. Line 2 in Algorithm 1. The used LP solver for the problem in (10) is the primal-dual interior point solver in [43].

[8]Parameters for Fig. 2 were $R_1 = R_2 = 42\,\text{Mbps}$, $R_3 = R_4 = 3\,\text{Mbps}$, $\hat{w}_1 = \hat{w}_2 = 0.5 * [0.05, 0.1, \ldots, 0.95]$, $\hat{w}_3 = \hat{w}_4 = 0.5 * [0.95, 0.9, \ldots, 0.05]$, $\breve{w}_u = 0, \forall u \in \mathcal{U}$, shorter and longer two lines have a length of 300 m and 1000 m, respectively, and are collocated at the CO side. Results for the upper three curves show non-dominated points only.

[9]The time-sharing problem in (8) is solved optimally by Algorithm 1, solving the subproblems in (12b) as stated in Lines 5 and 6 in Algorithm 1, and omitting Line 9 which only concerns the original problem in (2).

[10]For brevity we omitted results obtained under the continuous DSM scheme in [44], an energy-efficient modification of that in [45], and extended by CDBL as done for ISB to obtain discrete-rate solutions, which further support this statement.

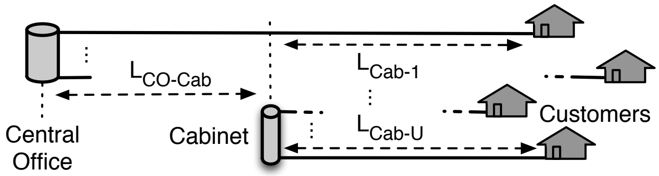

Fig. 3. Schematic of a mixed DSL deployment.

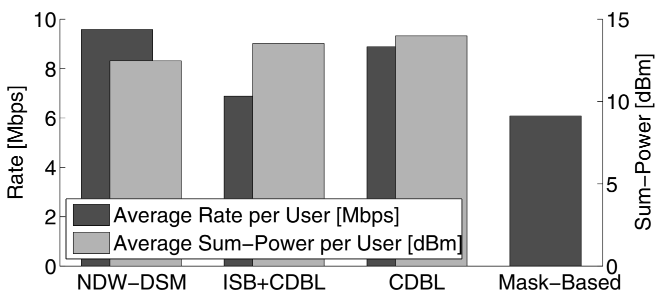

Fig. 4. Average per-user rate and sum-power in a 50 user VDSL scenario.[12]

LP solver. While we apply the interior-point solver in [43], a simplex solver [27] would result in a number of time-shared subcarriers as given in Theorem 2.

Fig. 2 also exemplifies the performance loss when the heuristic in [20] is used jointly with NDW-DSM, cf. Section III-C 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 in (2) of less than approximately 2 %. This is inferred by using the lower-bound on the optimal objective of (2) obtained by optimally solving the time-sharing relaxation in (8), cf. the lowest curve in Fig. 2. This shows the validity of our approach to optimize (2) by (e.g., approximately) solving the time-sharing problem in (8) and reconstructing solutions for (2) by the CHET heuristic.

B. Mixed DSL Example with 50 *Users*

Here we consider a 50 user upstream mixed VDSL scenario as shown in Fig. 3 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. Fig. 4 shows the results of various DSM schemes under rate-maximization ($\hat{\mathbf{w}} = \mathbf{0}, \check{\mathbf{w}} = \mathbf{1}$, cf. (3)) 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 a warm-start local search heuristic and a randomized heuristic based on SEGO, cf. [22] for a detailed description and parameters settings. The runtime of NDW-DSM in this scenario was around 6 hours, while that of ISB and CDBL 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" bit-loading, ISB and CDBL, respectively. Regarding solely the short lines NDW-DSM improves their sum-rate by 88.4 % compared to CDBL, demonstrating the disadvantage of long lines under the greedy DSM scheme CDBL. In terms of transmit sum-power NDW-DSM saves 21.6 % and 29.5 % compared to ISB and CDBL, respectively. In conclusion, this example demonstrates the possible gains of our combinatorial search compared to user-iterative and greedy DSM schemes under discrete rates.

C. Average Performance for Mixed 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 VDSL deployments as shown in Fig. 3 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 CDBL 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.

Fig. 5 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, CDBL and Mask-Based bit-loading, respectively. Regarding once more only the central-office deployed lines we have an average gain by NDW-DSM compared to CDBL of $+1.56 \pm 0.05$ Mbps (11.1 %), confirming the behavior of CDBL towards longer lines as observed in Section V-B.

VI. CONCLUSIONS

We proposed a novel power control algorithm for interference-limited discrete-rate multi-user and multi-carrier systems. It uses a specific stable dual decomposition scheme which allows for suboptimal combinatorial heuristics being used for the per-subcarrier subproblems, and a combination

[11]We adapted the mixing algorithm in [20] to the studied sum-power minimization problem by using the (normalized [20]) Euclidean distance to the optimal time-shared solution found by NDW-DSM as the selection metric.

[12]Sum-power results under the mask-based scheme were omitted as the target-rates could not be reached.

[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 [22] were parallelized in Matlab over 4 processes. We note that the exact runtimes may vary depending on the channel and network model. For example, in experiments using the measurements in [46] and assuming the lines are distributed in three binders with an inter-binder attenuation of 7.6 dB as reported in [47] we observed roughly 50 % higher simulation times for rate maximization under NDW-DSM and ISB, while that under CDBL is roughly proportional to the achieved bit-rate and was hence even three times as high. 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 II-B.

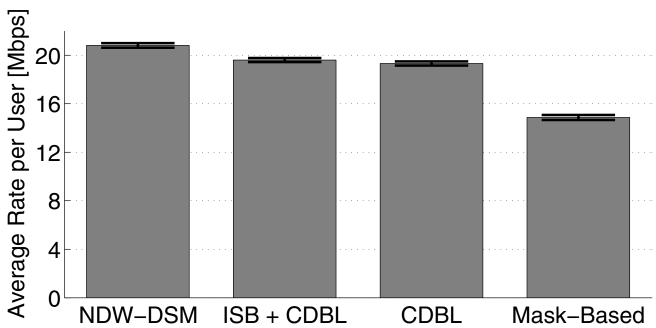

Fig. 5. Average rates in 10^3 DSL scenarios with 99 % confidence intervals.

heuristic to recover solutions to the original primal problem. An overview of the basic building blocks for various heuristics was given, including the greedy search, the local search, and a mixed exhaustive and greedy search procedure which exploits the presence of dominant crosstalkers. Performance gains using the novel power control framework using heuristics were quantified by simulations on specific near-far digital subscriber line (DSL) scenarios as well as a large set of thousand mixed DSL deployments. The results show an average gain in sum-rate by our algorithm of 6 % and 40 % compared to dynamic and static spectrum management schemes, respectively.

Appendix A
Proof of Theorem 2

Proof: We will analyze the number of feasible discrete power allocations $|\mathcal{Q}|$ per-subcarrier in an interference channel. The subcarrier indices are omitted 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 Δ. 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 Δ bits implies $r^u(\mathbf{p}) \geq \Delta, \forall u \in \mathcal{U}$. Reformulating (1) we obtain, $\forall u \in \mathcal{U}$,

$$1 \geq (2^{\Delta}-1)\Gamma\left(\sum\nolimits_{i\in\mathcal{U}\setminus u}\frac{H^{ui}}{H^{uu}}\frac{p^i}{p^u} + \frac{N^u}{H^{uu}p^u}\right). \quad (15)$$

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 (15) is given in (9). 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 [48, 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 $T_{\mathbf{r}}(k\cdot\Delta)$ of any number of bits $(k\cdot\Delta) \in \mathcal{B}, 0 \leq k \leq |\mathcal{B}|-1$, in $\mathbf{r}$. As any specific number of bits can only appear U times, we have for the number of types [48, Thm. 11.1.1]

$$|\mathcal{T}_U| \leq (U+1)^{|\mathcal{B}|} \triangleq m(U). \quad (16)$$

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

$$|\mathcal{S}(T)| = \binom{U}{UT(0),\ldots,UT(\hat{B})} \leq 2^{UH(T)} \quad (17a)$$

$$\leq 2^{U(\hat{B}/\Delta+1)\cdot\frac{\lceil U/(\hat{B}/\Delta+1)\rceil}{U}\cdot\log\left(\frac{\lceil U/(\hat{B}/\Delta+1)\rceil}{U}\right)} \triangleq n(U), \quad (17b)$$

where the first inequality follows from [48, 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 $T \in \mathcal{T}_{\hat{U}+t}$ given by

$$\mathcal{T}_{\hat{U}+t} = \left\{\left(\frac{\hat{U}\hat{T}(0)+t}{\hat{U}+t}, \frac{\hat{U}\hat{T}(\Delta)}{\hat{U}+t}, \ldots, \frac{\hat{U}\hat{T}(\hat{B})}{\hat{U}+t},\right) \mid \hat{T} \in \mathcal{T}_{\hat{U}}\right\}. \quad (18)$$

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 (18). Assuming any $U > \hat{U}+t, t > \hat{U}$, we have

$$|\mathcal{S}(T^{\hat{T}})| = \frac{U\cdot(U-1)\cdot\ldots\cdot(\hat{U}+1)}{(\hat{U}\hat{T}(0)+t)\cdot\ldots\cdot(\hat{U}\hat{T}(0)+1)}|\mathcal{S}(\hat{T})| \quad (19a)$$

$$\leq \frac{U\cdot(U-1)\cdot\ldots\cdot(\hat{U}+1)}{t\cdot(t-1)\cdot\ldots\cdot 1}n(\hat{U}) \quad (19b)$$

$$= \frac{U\cdot(U-1)\cdot\ldots\cdot(U-\hat{U}+1)}{\hat{U}!}n(\hat{U}) = O\left(U^{\hat{U}}\right) \quad (19c)$$

where in (19a) we use the fact that only the number of occurrences of 0 bits grows for $U \geq \hat{U}$, in (19b) we use the bound in (17b) and bound the expression by assuming $\hat{T}(0) = 0$, and in (19c) we use the assumption $t > \hat{U}$. Summarizing, for any $U > \hat{U}$ we have that $|\mathcal{T}_U| \leq m(\hat{U})$ and $|\mathcal{S}(T)|$ is polynomially bounded by (19c), $\forall T \in \mathcal{T}_U$, concluding the proof. ■

Appendix B
Proof of Theorem 3

Proof: Assume feasible weights $\tilde{\xi}', \tilde{\xi}_c^i, i \in \mathcal{I}_c^{(t)}, c \in \mathcal{C}$, in (10) with $0 < \tilde{\xi}' < 1$. Consider weights $\xi' = 0, \xi_c^i = \tilde{\xi}_c^i/(1-\tilde{\xi}'), \forall i \in \mathcal{I}_c^{(t)}$, for the problem in (10), setting the weights of columns not included in (10) to zero, i.e., $\xi_c^i = 0, \forall i \in \mathcal{I}_c \setminus \mathcal{I}_c^{(t)}, \forall c \in \mathcal{C}$. Regarding (10b)–(10d), from the feasibility of $\tilde{\xi}', \{\tilde{\xi}_c^i\}$ it follows that the weights $\xi', \{\xi_c^i\}$ are also feasible in (8) and (10). Using (3) we can write the objective value for weights $\xi', \{\xi_c^i\}$ in (10a) as

$$\sum_{c\in\mathcal{C}}\sum_{i\in\mathcal{I}_c^{(t)}}\left(\hat{\mathbf{w}}^T\mathbf{p}_c^i - \check{\mathbf{w}}^T\mathbf{r}_c(\mathbf{p}_c^i)\right)\xi_c^i \quad (20a)$$

$$\preceq \sum_{c\in\mathcal{C}}\sum_{i\in\mathcal{I}_c^{(t)}}\left(\hat{\mathbf{w}}^T\mathbf{p}_c^i - \check{\mathbf{w}}^T\mathbf{r}_c(\mathbf{p}_c^i)\right)\tilde{\xi}_c^i + \left(\hat{\mathbf{w}}^T\mathbf{P} - \check{\mathbf{w}}^T\mathbf{R}\right)\tilde{\xi}' \quad (20b)$$

$$\prec \sum_{c\in\mathcal{C}}\sum_{i\in\mathcal{I}_c^{(t)}}\left(\hat{\mathbf{w}}^T\mathbf{p}_c^i - \check{\mathbf{w}}^T\mathbf{r}_c(\mathbf{p}_c^i)\right)\tilde{\xi}_c^i + f'(\mathbf{P},\mathbf{R})\tilde{\xi}' \quad (20c)$$

where the two parts in (20b) are obtained by multiplying (20a) once by $(1-\tilde{\xi}')$ and once by $\tilde{\xi}'$, respectively, and using the definition of ξ_c^i above as well as feasibility in (10). Inequality (20c) follows as $\boldsymbol{\delta} \succ \mathbf{0}$. This implies that the weights $0 < \tilde{\xi}' < 1$ are not optimal in (10). Again from $\boldsymbol{\delta} \succ \mathbf{0}$ it follows that if weights as described in the theorem exist they will be feasible in (10) with $\tilde{\xi}' = 0$ and have a lower objective than $f'(\mathbf{P}, \mathbf{R})$, concluding $\tilde{\xi}' = 0$ at optimum of (10). On the other hand, if no weights as described in the theorem exist we must have $\xi' > 0$ at optimum of (10), $0 < \xi' < 1$ would lead to a contradiction by the arguments above, and hence $\xi' = 1$. ■

REFERENCES

[1] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.

[2] M. Barton and M. Honig, "Spectral optimization of discrete multitone system on twisted wire copper plants," in *Proc. 1995 IEEE Int. Conf. Commun.*, vol. 3, pp. 1668–1672.

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

[4] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, July 2006.

[5] R. Cendrillon and M. Moonen, "Iterative spectrum balancing for digital subscriber lines," in *Proc. 2005 IEEE Int. Conf. Commun.*,, vol. 3, pp. 1937–1941.

[6] P. Tsiaflakis, J. Vangorp, M. Moonen, and J. Verlinden, "A low complexity optimal spectrum balancing algorithm for digital subscriber lines," *Signal Process.*, vol. 87, no. 7, pp. 1735–1753, July 2007.

[7] J. Papandriopoulos and J. S. Evans, "Low-complexity distributed algorithms for spectrum balancing in multi-user DSL networks," in *Proc. 2006 IEEE Int. Conf. Commun.*, vol. 7, pp. 3270–3275.

[8] P. Tsiaflakis, I. Necoara, J. Suykens, and M. Moonen, "Improved dual decomposition based optimization for DSL dynamic spectrum management," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2230–2245, Apr. 2010.

[9] S. Huberman, C. Leung, and T. Le-Ngoc, "Dynamic spectrum management (DSM) algorithms for multi-user xDSL," *IEEE Commun. Surveys Tutorials*, vol. 14, no. 1, pp. 109–130, first quarter 2012.

[10] Z.-Q. Luo and S. Zhang, "Duality gap estimation and polynomial time approximation for optimal spectrum management," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2675–2689, July 2009.

[11] C. Wong, R. Cheng, K. Letaief, and R. Murch, "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758, Oct. 1999.

[12] W. Rhee and J. Cioffi, "Increase in capacity of multiuser OFDM system using dynamic subchannel allocation," in *Proc. 2000 IEEE Veh. Technol. Conf. – Spring*, vol. 2, pp. 1085–1089.

[13] M. Wolkerstorfer, T. Nordström, B. Krasniqi, M. Wrulich, and C. Mecklenbräuker, "OFDM/OFDMA subcarrier allocation," in *Cross Llayer Designs in WLAN Systems*, N. Zorba, C. Skianis, and C. Verikoukis, editors. Troubador Publishing Ltd., 2011, vol. 1, ch. 6, pp. 177–216.

[14] A. Garg and A. Chaturvedi, "Generalization of hybrid time divisioning for power allocation in DMT-based DSL systems," *IEEE Commun. Lett.*, vol. 11, no. 6, pp. 504–506, June 2007.

[15] M. Wolkerstorfer, D. Statovci, and T. Nordström, "Duality-gap bounds for multi-carrier systems and their application to periodic scheduling," in *Proc. 2010 IEEE Int. Conf. Commun.*

[16] J. Lee, R. Sonalkar, and J. Cioffi, "Multi-user discrete bit-loading for DMT-based DSL systems," in *Proc. 2002 IEEE Global Commun. Conf.*, vol. 2, pp. 1259–1263.

[17] B. Krongold, "Optimal efficient discrete resource allocation for OFDMA systems," in *Proc. 2008 IEEE Wireless Commun. Netw. Conf.*, pp. 1722–1727.

[18] M. Patriksson, "Simplicial decomposition algorithms," in *Encyclopedia of Optimization*, 2nd edition, C. Floudas and P. Pardalos, editors. Springer, 2009, pp. 3579–3585.

[19] A. Forouzan, "Optimal spectrum management of DSL with nonstrictly convex rate region," *IEEE Trans. Signal Process.*, vol. 57, no. 7, pp. 2558–2568, July 2009.

[20] A. Forouzan and M. Moonen, "Lagrange multiplier optimization for optimal spectrum balancing of DSL with logarithmic complexity," in *Proc. 2011 IEEE Int. Conf. Commun.*

[21] R. Lui and W. Yu, "Low-complexity near-optimal spectrum balancing for digital subscriber lines," in *Proc. 2005 IEEE Int. Conf. Commun.*, vol. 3, pp. 1947–1951.

[22] M. Wolkerstorfer and T. Nordström, "Coverage optimization in DSL networks by low-complexity discrete spectrum balancing," in *Proc. 2011 IEEE Global Commun. Conf.*

[23] ——, "Heuristics for discrete power control—a case-study in multi-carrier DSL networks," in *Proc. 2011 ALIO/EURO Workshop Applied Combinatorial Optimization*.

[24] P. Golden, H. Dedieu, and K. Jacobsen, editors, *Fundamentals of DSL Technology*. Auerbach Publications, 2006.

[25] R. Yates, "A framework for uplink power control in cellular radio systems," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 7, pp. 1341–1347, Sep. 1995.

[26] Y. Nakagawa, "An improved surrogate constraints method for separable nonlinear integer programming," *J. Operations Research*, vol. 46, no. 2, pp. 145–163, 2003.

[27] D. Bertsimas and J. N. Tsitsiklis, *Introduction to Linear Optimization*. Athena Scientific, 1997.

[28] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, 1982.

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

[30] M. Lübbecke and J. Desrosiers, "Selected topics in column generation," *Operations Research*, vol. 53, no. 6, pp. 1007–1023, Nov.–Dec. 2005.

[31] G. Dantzig, *Linear Programming and Extensions*. Princeton University Press, 1963.

[32] P. Björklund, P. Värbrand, and D. Yuan, "Resource optimization of spatial TDMA in ad hoc radio networks: a column generation approach," in *Proc. 2003 IEEE Int. Conf. Computer Commun.*,, vol. 2, pp. 818–824.

[33] M. Johansson and L. Xiao, "Cross-layer optimization of wireless networks using nonlinear column generation," *IEEE Trans. Wireless Commun.*, vol. 5, no. 2, pp. 435–445, Feb. 2006.

[34] D. Barreto and S. Chiu, "Decomposition methods for cross-layer optimization in wireless networks," in *Proc. 2007 IEEE Wireless Commun. Netw. Conf.*, pp. 270–275.

[35] P. Soldati and M. Johansson, "Dynamic resource allocation in OFDMA multi-cellular systems," School of Electrical Engineering (KTH), Technical Report TRITA-EE 2008:048, Aug. 2008.

[36] M. Wolkerstorfer, J. Jaldén, and T. Nordström, "Low-complexity optimal discrete-rate spectrum balancing in digital subscriber lines," to appear in *Elsevier Signal Process.*, 2012.

[37] R. Srikant, *The Mathematics of Internet Congestion Control*. Birkhäuser, 2004.

[38] M. Wolkerstorfer, "Energy-efficient resource allocation in multi-carrier digital subscriber lines," Ph.D. dissertation, Vienna University of Technology, Vienna, Austria, Feb. 2012.

[39] ETSI, "Transmission and multiplexing (TM); access transmission systems on metallic access cables; very high speed Digital Subscriber Line (VDSL); part 1: functional requirements," ETSI, Tech. Rep. TM6 TS 101 270-1, Version 1.3.1, July 2003.

[40] T. Nordström, *et.al.*, "xDSL simulator v3.1," xdsl.ftw.at, Oct. 2008.

[41] J. Campello, "Optimal discrete bit loading for multicarrier modulation systems," in *Proc. 1998 IEEE Int. Symp. Inf. Theory*, p. 193.

[42] D. Yu, K. Seong, and J. Cioffi, "Multiuser discrete bit-loading for digital subscriber lines," in *Proc. 2007 IEEE Int. Conf. Commun.*, pp. 2755–2760.

[43] The MOSEK Optimization Toolbox for MATLAB Manual, Version 6.0 (Revision 106), MOSEK ApS. Denmark, 2011.

[44] M. Wolkerstorfer, D. Statovci, and T. Nordström, "Dynamic spectrum management for energy-efficient transmission in DSL," in *Proc. 2008 IEEE Int. Conf. Commun. Syst.*

[45] J. Papandriopoulos and J. Evans, "SCALE: a low-complexity distributed protocol for spectrum balancing in multiuser DSL networks," *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711–3724, Aug. 2009.

[46] D. Statovci and T. Nordström, "Adaptive resource allocation in multiuser FDD-DMT systems," in *Proc. 2004 European Signal Process. Conf.*, pp. 1213–1216.

[47] J. Maes, M. Guenach, and M. Peeters, "Statistical channel model for gain quantification of DSL crosstalk mitigation techniques," in *Proc. 2009 IEEE Int. Conf. Commun.*, pp. 1–5.

[48] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd edition. Wiley Interscience, 2006.

Martin Wolkerstorfer (S'09-M'12) received the "Dipl. Ing." degree (equivalent to a master's degree) in electrical engineering from Graz University of Technology, Austria, in 2007, and the Ph.D. degree from Vienna University of Technology, Austria, in 2012, respectively. He is currently working as a senior researcher in the field of signal and information processing at the FTW Telecommunications Research Center Vienna, Austria. His research interests include the application of optimization theory in communications and signal processing, and the energy-efficient operation of broadband access networks such as DSL and WLAN.

Joakim Jaldén (S'03-M'08) received the M.Sc. and Ph.D. in electrical engineering from the Royal Institute of Technology (KTH), Stockholm, Sweden, in 2002 and 2007, respectively. From July 2007 to June 2009, he held a post-doctoral research position at the Vienna University of Technology, Vienna, Austria. He also studied at Stanford University, CA, USA, from September 2000 to May 2002, and worked at ETH, Zürich, Switzerland, as a visiting researcher, from August to September 2008. In July 2009, he joined the Signal Processing Lab within the School of Electrical Engineering at KTH, Stockholm, Sweden, as an Assistant Professor. He served as an associate editor for IEEE COMMUNICATIONS LETTERS between 2009 and 2011 and has been serving as an associate editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING since 2012.

For his work on MIMO communications, Joakim has been awarded the IEEE Signal Processing Society's 2006 Young Author Best Paper Award and the first prize in the Student Paper Contest at the 2007 International Conference on Acoustics, Speech and Signal Processing (ICASSP). He is also a recipient of the Ingvar Carlsson Award issued in 2009 by the Swedish Foundation for Strategic Research.

Tomas Nordström (S'88-A'95-M'00-SM'01) received the M.S.E.E. degree in 1988, the licentiate degree in 1991, and the Ph.D. degree in 1995, all from Luleå University of Technology, Sweden. Between 1996 and 1999, he was with Telia Research (the research branch of the Swedish incumbent telephone operator) where he developed broadband Internet communication over twisted copper pairs. In December 1999, he joined the FTW Telecommunications Research Center Vienna, Austria, where he currently works as a Key Researcher in the field of "broadband wireline access." During his years at FTW, he worked on various aspects of wireline communications such as the simulation of xDSL systems, cable characterization, RFI suppression, exploiting the common-mode signal in xDSL, and more recently, dynamic spectrum management. Since 2009, he has held a position as an Associate Professor in computer systems engineering at Halmstad University (HH), Sweden. His current research interests include all aspects of energy efficient computer and communication systems, from cross-layer design and signal processing all the way to novel power amplifier designs and multi-core computing.