Network Utility Maximization for Adaptive Resource Allocation in DSL Systems

Jeroen Verdyck
STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics
ESAT Department of Electrical Engineering
KU Leuven
Leuven, Belgium
jeroen.verdyck@esat.kuleuven.be

Chris Blondia
IDLab research group
Dept. Mathematics and Computer Science
University of Antwerp-imec
Antwerp, Belgium
chris.blondia@uantwerpen.be

Marc Moonen
STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics
ESAT Department of Electrical Engineering
KU Leuven
Leuven, Belgium
marc.moonen@esat.kuleuven.be

***Abstract*—When signal coordination techniques can not eliminate all crosstalk in a digital subscriber line (DSL) system, competition for data rate among different users is strong. In such scenarios, employing a static resource allocation fails to capitalize on the time dependent nature of the traffic carried by the DSL network. An alternative approach is adaptive resource allocation, consisting of dividing time into slots of short duration and using a different resource allocation in each slot. A cross-layer scheduler then decides on the resource allocation for each time slot by solving a network utility maximization (NUM) problem. For many DSL systems however, this NUM problem is non-convex and solving it is NP-hard. This paper presents a fast algorithm for finding a local solution to the NUM problem, which is referred to as NUM-DSB. The algorithm is able to handle many DSL deployment scenarios, and is applicable regardless of the utility function's properties.**

***Index Terms*—DSL, Crosstalk, Cross layer design, Adaptive resource allocation, Minorize-maximization**

I. Introduction

Dynamic spectrum management (DSM) techniques, which are used in digital subscriber line (DSL) systems to fight crosstalk, are often divided into two categories: spectrum coordination and vectoring. Spectrum coordination reduces crosstalk by jointly managing the transmit power spectra of all users, whereas vectoring encompasses jointly coordinating multiple lines on a signal level. As vectoring techniques are often able to fully eliminate the effects of crosstalk, they have become the industry standard as of VDSL2 [1].

In many practical deployment scenarios however, the elimination of all crosstalk through vectoring is inhibited, such that one has to rely on spectrum coordination to tone down the effects of the residual crosstalk. Examples include fiber to the frontage deployments implementing multi-user full-duplex signaling [2], [3], and deployments where multiple DSL access multiplexers (DSLAMs) are active on a single cable binder [4]. Optimizing the performance in such a system corresponds to solving a multi-objective optimization problem, yielding a collection of Pareto optimal resource allocation configurations from which one has to choose a single configuration for transmission. Often, a conservative approach is followed where the physical layer is configured statically over time using a single Pareto optimal resource allocation.

An alternative to this static resource allocation is adaptive resource allocation (ARA), which consists of dividing time into slots of short duration, and changing the resource allocation from one time slot to the next. A cross-layer scheduler then chooses one setting for each time slot in accordance with upper layer requirements. To this end, the cross-layer scheduler defines a utility function U, and solves the resulting network utility maximization (NUM) problem, i.e.

$$\underset{\boldsymbol{x}\in\mathcal{X}}{\text{maximize}}\, U([r_1(\boldsymbol{x}), ..., r_N(\boldsymbol{x})]^{\mathrm{T}}), \tag{1}$$

where $\boldsymbol{x}$ is the physical layer resource allocation, and $r_n(\boldsymbol{x})$ denotes the resulting data rate for user n. Algorithms finding the global optimum of problem (1) are available in [5], [6]. It has been shown that adaptive resource allocation can lead to significantly improved throughput and delay performance [7], [8], as well as to a reduced power consumption [9].

As the rate region of a DSL system is a convex set [10], the solution to any NUM problem can be found by solving a sequence of weighted sum rate (WSR) maximization subproblems [5], as in Algorithm 1. Convergence of such WSR based NUM algorithms however requires finding the global optimum of each WSR maximization subproblem. As finding the globally optimal solution to the WSR maximization subproblem yielded by many DSM techniques, including the techniques considered in this paper, is itself NP-hard [11], directly applying Algorithm 1 to problem (1) results in an exceedingly high computational complexity. This is especially problematic in the context of ARA, where a large number of NUM problems is to be solved.

This paper therefore proposes an new algorithm for the NUM problem (1), which will be referred to as NUM-DSB. The algorithm is highly generic - both in the sense of being

This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of 1) Fonds de la Recherche Scientifique - FNRS and the Fonds Wetenschappelijk Onderzoek - Vlaanderen under EOS Project no 30452698 '(MUSE-WINET) MUlti-SErvice WIreless NETwork', 2) Research Project FWO nr. G.0B1818N 'Real-time adaptive cross-layer dynamic spectrum management for fifth generation broadband copper access networks', 3) VLAIO O&O Project nr. HBC.2016.0055 '5GBB Fifth generation broadband access'. The scientific responsibility is assumed by its authors.

able to handle many deployment scenarios, and of being applicable regardless of the utility function's properties. NUM-DSB is based on distributed spectrum balancing (DSB) [12] for WSR maximization in systems that only employ spectrum coordination.

Algorithm 1 Sequential WSR maximization for NUM

Initialize $\boldsymbol{\omega} \in \mathbb{R}_{+}^{N}$
for i = 0, 1, ... **do**
  Solve $\text{maximize}_{\boldsymbol{x} \in \mathcal{X}}\ \boldsymbol{\omega}^{\mathrm{T}} \boldsymbol{r}(\boldsymbol{x})$
  Update $\boldsymbol{\omega}$

II. DSB AS A MINORIZE-MAXIMIZATION ALGORITHM

DSB [12], [13] is a low-complexity algorithm for finding a local solution of the WSR maximization problem in DSL systems that do not employ any signal coordination. Initially, DSB has been introduced as a fixed point algorithm aimed at finding a solution to the KKT stationarity conditions [12]. In this section, a generalized version of DSB is described, encompassing coupling constraints between users. In the exposition, DSB is interpreted as a minorize-maximization (MM) algorithm [14]. This interpretation is in line with recent optimization literature [15]–[17], and more easily lends itself to be extended to NUM problems.

DSB can be applied to optimization problems of the form

$$\underset{\boldsymbol{x}}{\text{maximize}} \quad \sum_{i \in \mathcal{I}} \sum_{n \in \mathcal{N}_i} \omega_{i,n} r_{i,n}(\boldsymbol{x}) \tag{2a}$$

$$\text{subject to} \quad \boldsymbol{x}_i \in \mathcal{X}_i,\ \forall i \in \mathcal{I}, \tag{2b}$$

where $\mathcal{I} \triangleq \{1, ..., I\}$ is a set of group indices, and where $\mathcal{N}_i \triangleq \{1, \ldots, N_i\}$ denotes the set of users that are in group i. The decision variable $\boldsymbol{x}$ consists of multiple coordinate blocks, i.e. $\boldsymbol{x}^{\mathrm{T}} = [\boldsymbol{x}_1{}^{\mathrm{T}}, ..., \boldsymbol{x}_I{}^{\mathrm{T}}]$, where coordinate block $\boldsymbol{x}_i$ corresponds to the resource allocation variables of the users in group i. Moreover, each rate function $r_{i,n}$ is assumed to be concave in the coordinate block associated with its group $\boldsymbol{x}_i$, and convex in all other coordinate blocks $\boldsymbol{x}_j$, where $j \in \mathcal{I} \setminus i$.

Many DSM techniques yield resource allocation problems that fit problem (2). Examples include the power allocation problem in DSL systems employing only spectrum coordination [12], the joint transmit power allocation and receive filter design problem in upstream DSL systems [13], the power allocation problem in multi-user full-duplex DSL systems employing zero-forcing precoders and receive filters [2], and the power allocation problem in downstream grouped vectoring (GV) scenarios with zero-forcing precoders [4]. This last problem will be studied more extensively in Section IV.

In each iteration, DSB solves a set of surrogate problems, one for each group $i \in \mathcal{I}$, which are much easier to solve than the original WSR maximization problem (2). The surrogate problem of group i is obtained from (2) by restricting the maximization to $\boldsymbol{x}_i$, i.e. considering all variables $\boldsymbol{x}_j$ with $j \in \mathcal{I} \setminus i$ to be constant, and approximating the rate of the users not in $\mathcal{N}_i$ with a first order Taylor expansion, i.e.

$$\underset{\boldsymbol{x}_i}{\text{maximize}} \quad \sum_{j \in \mathcal{I}} \sum_{n \in \mathcal{N}_j} \omega_{j,n} \tilde{r}_{j,n}^{(i)}(\boldsymbol{x}_i; \bar{\boldsymbol{x}}) \tag{3a}$$

$$\text{subject to} \quad \boldsymbol{x}_i \in \mathcal{X}_i. \tag{3b}$$

In (3), $\bar{\boldsymbol{x}}$ is the current value of $\boldsymbol{x}$, and $\tilde{r}_{j,n}^{(i)}$ is defined as

$$\tilde{r}_{j,n}^{(i)}(\boldsymbol{x}_i; \bar{\boldsymbol{x}}) \triangleq \begin{cases} r_{j,n}([\bar{\boldsymbol{x}}_1^{\mathrm{T}}, ..., \boldsymbol{x}_i{}^{\mathrm{T}}, ..., \bar{\boldsymbol{x}}_I^{\mathrm{T}}]^{\mathrm{T}}) & \text{if } j = i \\ \bar{r}_{j,n} + (\boldsymbol{x}_i - \bar{\boldsymbol{x}}_i)^{\mathrm{T}} \nabla_{\boldsymbol{x}_i} r_{j,n}\big|_{\boldsymbol{x}=\bar{\boldsymbol{x}}} & \text{if } j \neq i \end{cases} \tag{4}$$

where $\bar{r}_{j,n} \triangleq r_{j,n}(\bar{\boldsymbol{x}})$. After solving a set of surrogate problems, either a single or all coordinate blocks of $\bar{\boldsymbol{x}}$ are updated (these modes of operation are respectively referred to as 'Gauss-Seidel' (GS) and 'Jacobi' updating), and a new set of surrogate problems is constructed. The resulting sequence of values for $\bar{\boldsymbol{x}}$ *can* converge to a stationary point of the original optimization problem (2). Moreover, each iterate $\bar{\boldsymbol{x}}$ generated by the algorithm is feasible, such that DSB complies with the definition of a real-time DSM algorithm [18].

Due to its convexity in $\boldsymbol{x}_i$ for $j \neq i$, $\tilde{r}_{j,n}^{(i)}$ constitutes a lower bound approximation on $r_{j,n}$ which is exact in $\bar{\boldsymbol{x}}$, i.e.

$$r_{j,n}([\bar{\boldsymbol{x}}_1^{\mathrm{T}}, ..., \boldsymbol{x}_i{}^{\mathrm{T}}, ..., \bar{\boldsymbol{x}}_I^{\mathrm{T}}]^{\mathrm{T}}) \geq \tilde{r}_{j,n}^{(i)}(\boldsymbol{x}_i; \bar{\boldsymbol{x}}) \tag{5a}$$

$$r_{j,n}(\bar{\boldsymbol{x}}) = \tilde{r}_{j,n}^{(i)}(\bar{\boldsymbol{x}}_i; \bar{\boldsymbol{x}}). \tag{5b}$$

Similarly, the objective in (3a) is a lower bound on the objective in (2a). When employing GS updating, DSB can thus be interpreted as a block-wise (MM) algorithm which is guaranteed to converge when as the objective of (3a) is bounded from above. As the surrogate functions (4) also satisfy [15, Assumption 2], it can additionally be proven that DSB with GS updating indeed converges to a stationary point of problem (2).

It is noted that in this section, DSL nomenclature has been adopted by referring to the algorithm as 'DSB'. The general algorithmic structure is however known as 'successive convex approximation' (SCA). SCA algorithms with proven convergence that include DSB as a special case, are BSUM [15], SJBR [16] and FLEXA [17].

III. NUM-DSB

This section shows how the DSB algorithm can be extended such that it can be applied to NUM problems. The resulting algorithm will be referred to as DSB for network utility maximization (NUM-DSB), and can be applied to NUM problems of the following form.

$$\underset{\boldsymbol{x}}{\text{maximize}} \quad U\left(\boldsymbol{r}(\boldsymbol{x})\right) \tag{6a}$$

$$\text{subject to} \quad \boldsymbol{x}_i \in \mathcal{X}_i,\ \forall i \in \mathcal{I} \tag{6b}$$

In (6), $\boldsymbol{r}(\boldsymbol{x})$ contains the rate vectors of all groups, i.e. $\boldsymbol{r}(\boldsymbol{x})^{\mathrm{T}} = [\boldsymbol{r}_1(\boldsymbol{x})^{\mathrm{T}}, ..., \boldsymbol{r}_I(\boldsymbol{x})^{\mathrm{T}}]$. In turn, $\boldsymbol{r}_i$ contains the rate of all users in group i, i.e. $\boldsymbol{r}_I(\boldsymbol{x})^{\mathrm{T}} = [r_{i,1}(\boldsymbol{x}), ..., r_{i,N_i}(\boldsymbol{x})]$. As before, each rate function $r_{i,n}(\boldsymbol{x})$ is assumed to be concave in its associated coordinate block $\boldsymbol{x}_i$, and convex in all other coordinate blocks $\boldsymbol{x}_j$, where $j \in \mathcal{I} \setminus i$. It should be noted that the same DSM

ISBN 978-90-827970-1-5 © EURASIP 2018

techniques that yield WSR maximization problems fitting (2), will also yield NUM problems that fit (6).

In each iteration, NUM-DSB solves a set of surrogate problems, one for each group $i \in \mathcal{I}$. The surrogate problem of group i is obtained from problem (1) by, again, restricting the maximization to $\boldsymbol{x}_i$, and approximating the rate of all users not in $\mathcal{N}_i$ with a first order Taylor expansion.

$$\begin{aligned} &\underset{\boldsymbol{x}_i}{\text{maximize}} && U\left(\tilde{\boldsymbol{r}}^{(i)}(\boldsymbol{x}_i;\bar{\boldsymbol{x}})\right) && \text{(7a)} \\ &\text{subject to} && \boldsymbol{x}_i \in \mathcal{X}_i && \text{(7b)} \\ & && \tilde{r}^{(i)}_{j,n}(\boldsymbol{x}_i;\bar{\boldsymbol{x}}) \geq 0,\ \forall j \in \mathcal{I}\setminus\{i\} && \text{(7c)} \end{aligned}$$

In (7), $\tilde{\boldsymbol{r}}^{(i)}_{j,n}$ is defined as in (4). As the utility function U might be undefined for negative rates, the NUM-DSB surrogate problem enforces positivity of U's arguments by including a positivity constraint on the approximate rates (7c). After solving (7), $\bar{\boldsymbol{x}}$ is updated and a new set of surrogate problems is constructed. As before, each iterate $\bar{\boldsymbol{x}}$ that is generated by NUM-DSB is feasible, such that NUM-DSB complies with the definition of a real-time DSM algorithm [18].

Each NUM-DSB surrogate problem (7) is solved using Algorithm 1. In turn, each iteration of Algorithm 1 requires solving a problem that is, apart from positivity constraint (7c), identical to the DSB surrogate problem (3). Assuming that the DSB surrogate problem can be solved efficiently, it is likely that each WSR maximization subproblem obtained by applying Algorithm 1 to problem (7) can be as well.

Combining the lower bounding property of (5a)-(5b) with the assumption that U is monotonically increasing in all its components, NUM-DSB with GS updating can still be interpreted as a block-wise MM algorithm. Moreover, as the objective of (6a) is bounded from above, NUM-DSB is guaranteed to converge regardless of the utility function's properties. No such convergence guarantees can be given when Jacobi updating is employed. However, NUM-DSB with Jacobi updating has been seen to converge in all numerical experiments that have been executed.

IV. Application: downstream grouped vectoring

This section considers a downstream grouped vectoring scenario. Each index $i \in \mathcal{I}$ now corresponds to a single vectoring group that contains users $\mathcal{N}_i$. After introducing the system model for such DSL systems, the exposition will focus on showing that the WSR problems yielded by applying Algorithm 1 to problem (7) can be solved efficiently.

DSL systems employ discrete multi-tone (DMT) modulation to split the available spectrum into a set of K tones, which is denoted as $\mathcal{K} = \{1, ..., K\}$. It is assumed that there is no so-called inter-carrier interference and hence transmission can be modeled on each tone k independently as

$$\begin{bmatrix} \boldsymbol{y}_{k,1} \\ \vdots \\ \boldsymbol{y}_{k,I} \end{bmatrix} = \begin{bmatrix} \boldsymbol{H}_{k,11} & \cdots & \boldsymbol{H}_{k,1I} \\ \vdots & \ddots & \vdots \\ \boldsymbol{H}_{k,I1} & \cdots & \boldsymbol{H}_{k,II} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{x}}_{k,1} \\ \vdots \\ \hat{\boldsymbol{x}}_{k,I} \end{bmatrix} + \begin{bmatrix} \boldsymbol{z}_{k,1} \\ \vdots \\ \boldsymbol{z}_{k,I} \end{bmatrix}, \quad (8)$$

where $\hat{\boldsymbol{x}}_{k,i}$ is an $N_i \times 1$ vector containing the signals transmitted by the i-th vectoring group on tone k. Similarly, $\boldsymbol{y}_{k,i}$ and $\boldsymbol{z}_{k,i}$ respectively contain the received signals and additive Gaussian noise of the i-th vectoring group on tone k. Moreover, $\boldsymbol{H}_{k,ij}$ is an $N_i \times N_j$ channel matrix with $[\boldsymbol{H}_{k,ij}]_{nm} = h_{k,ij,nm}$ the transfer function between transmitter m of group j and receiver n of group i, evaluated on tone k.

Signal coordination among users in the same group is possible at the transmitter side. Transmitted signals are generated as

$$\hat{\boldsymbol{x}}_{k,i} = \boldsymbol{T}_{k,i}\boldsymbol{x}_{k,i} \quad (9)$$

in group i, where $\boldsymbol{T}_{k,i}$ is the precoding matrix of group i on tone k, and where $\boldsymbol{x}_{k,i}{}^{\mathrm{T}} = [x_{k,i,1}, ..., x_{k,i,N_i}]$ contains the transmitted symbols of all users n in group i on tone k. Furthermore, the symbol power of user n in group i on tone k is defined as $s_{k,i,n} = \Delta_f \mathbb{E}[|x_{k,i,n}|^2]$, with Δ_f the tone spacing. The aggregate transmit power (ATP) on the n-th line of group i is then calculated as

$$P_{i,n} = \sum_{k\in\mathcal{K}} [\boldsymbol{T}_{k,i}\,\mathrm{diag}\,(\boldsymbol{s}_{k,i})\boldsymbol{T}_{k,i}{}^{\mathcal{H}}]_{nn}, \quad (10)$$

with $\boldsymbol{s}_{k,i}{}^{\mathrm{T}} = [s_{k,i,1}, ..., s_{k,i,N_i}]$. In G.fast, the ATP of each line is limited to $P_{\mathrm{tot}} = 4\,\mathrm{dBm}$.

It is assumed that each vectoring group $i \in \mathcal{I}$ uses a zero-forcing precoder, i.e.

$$\boldsymbol{T}_{k,i} = \boldsymbol{H}_{k,ii}{}^{-1}. \quad (11)$$

In order to simplify notation, the effective channel matrix $\hat{\boldsymbol{H}}_{k,ji} = \boldsymbol{H}_{k,ji}\boldsymbol{T}_{k,i}$ is introduced. The resulting SINR of user n in group i on tone k is then

$$\gamma_{k,i,n} = \frac{s_{k,i,n}}{\sigma_{k,i,n} + \sum_{j\in\mathcal{I}\setminus\{i\}}\sum_{m=1}^{N_j} |\hat{h}_{k,ij,nm}|^2 s_{k,j,m}}, \quad (12)$$

with $\sigma_{k,i,n} = \Delta_f \mathbb{E}[|z_{k,i,n}|]$ the received noise power of user n in group i on tone k. When the number of interferers $\sum_{j\in\mathcal{I}\setminus\{i\}} N_j$ is large, then the interference-plus-noise received by the users of group i is well approximated by a Gaussian distribution. Under this assumption, bitloading of user n in group i on tone k is accurately modeled by

$$b_{k,i,n} = \log_2(1 + \Gamma^{-1}\gamma_{k,i,n}), \quad (13)$$

where Γ denotes the SNR gap to capacity. The rate of user n in group i is then

$$r_{i,n} = f_{\mathrm{s}} \sum_{k\in\mathcal{K}} b_{k,i,n}, \quad (14)$$

with f_{s} the DMT symbol rate.

The above DSL system fits problem (6), regardless of the considered utility function. The different groups in (6) are given by the vectoring groups in $\mathcal{I}$, and where the feasible sets are defined as

$$\mathcal{X}_i = \{\boldsymbol{s}_i \in \mathbb{R}_+^{KN_i} \mid P_{i,n} \leq P_{\mathrm{tot}},\ \forall n \in \mathcal{N}_i\}, \quad (15)$$

with $\boldsymbol{s}_i = [s_{1,i,1}, ..., s_{K,i,1}, s_{1,i,2}, ..., s_{K,i,N_i}]^{\mathrm{T}}$. It is now shown that the WSR subproblem, obtained by applying Algorithm 1 to the NUM problem in (7), can be solved efficiently for the considered DSL system.

First, Lagrange dual decomposition is applied to decouple the WSR problem over users and tones. The Lagrangian $\mathcal{L}_i$ is obtained by dualizing the ATP and rate positivity constraints, i.e.

$$\mathcal{L}_i = \sum_{j\in\mathcal{I}} \sum_{n\in\mathcal{N}_j} \omega_{j,n} \tilde{r}_{j,n}^{(i)}(\boldsymbol{s}_i;\bar{\boldsymbol{s}}) + \sum_{j\in\mathcal{I}\setminus\{i\}} \sum_{n\in\mathcal{N}_j} \lambda_{j,n}^{\mathrm{r}} \tilde{r}_{j,n}^{(i)}(\boldsymbol{s}_i;\bar{\boldsymbol{s}}) + \sum_{n\in\mathcal{N}_i} \lambda_{i,n}^{\mathrm{p}} P_{\mathrm{tot}} - \sum_{n\in\mathcal{N}_i} \sum_{k\in\mathcal{K}} \boldsymbol{t}_{k,i,n}{}^{\mathcal{H}} \boldsymbol{\Lambda}^{\mathrm{p}} \boldsymbol{t}_{k,i,n} s_{k,i,n}, \quad (16)$$

with $\boldsymbol{t}_{k,i,n}$ the n-th column of $\boldsymbol{T}_{k,i}$, and where $\boldsymbol{\lambda}^{\mathrm{p}}$ and $\boldsymbol{\lambda}^{\mathrm{r}}$ respectively contain the Lagrange multipliers associated with the ATP and rate positivity constraints. The resulting Lagrange dual function is then given by

$$g_i(\boldsymbol{\lambda}^{\mathrm{p}}, \boldsymbol{\lambda}^{\mathrm{r}}) = \sup_{\boldsymbol{s}_i \in \mathbb{R}_+^{K\times N_i}} \{\mathcal{L}_i\}. \quad (17)$$

As the considered WSR subproblem is convex, the duality gap with problem (17) is zero. The maximization in (17) is separable per user and per tone, and has a unique maximizer that is given by (18) at the top of the next page.

It can be proven that the dual function as defined in (17) is continuously differentiable on its domain, with its partial derivatives given by

$$\frac{\partial g(\boldsymbol{\lambda}^{\mathrm{p}}, \boldsymbol{\lambda}^{\mathrm{r}})}{\partial \lambda_{i,n}^{\mathrm{p}}} = P_{\mathrm{tot}} \mathbf{1} - \sum_{k\in\mathcal{K}} [\boldsymbol{T}_{k,i} \operatorname{diag}(\boldsymbol{s}_{k,i}^{\star}) \boldsymbol{T}_{k,i}{}^{\mathcal{H}}]_{n,n}, \quad (19)$$

$$\frac{\partial g(\boldsymbol{\lambda}^{\mathrm{p}}, \boldsymbol{\lambda}^{\mathrm{r}})}{\partial \lambda_{j,n}^{\mathrm{r}}} = \tilde{r}_{j,n}^{(i)}(\boldsymbol{s}_i^{\star};\bar{\boldsymbol{s}}). \quad (20)$$

As the dual function is differentiable, one can employ gradient descent methods to solve the Lagrange dual problem, which is given by

$$\underset{\boldsymbol{\lambda}^{\mathrm{p}}\geq 0, \boldsymbol{\lambda}^{\mathrm{r}}\geq 0}{\text{minimize}} \quad g_i(\boldsymbol{\lambda}^{\mathrm{p}}, \boldsymbol{\lambda}^{\mathrm{r}}). \quad (21)$$

Differentiability of g_i can be established using results from [19]. First, define function $F_{i,n}$ and its convex conjugate $F_{i,n}^*$ as in [19, Chapter 11], i.e.

$$F_{i,n}(\boldsymbol{s}_{i,n}) = \begin{cases} -\tilde{r}_{i,n}^{(i)}([\mathbf{0}^{\mathrm{T}},...,\boldsymbol{s}_{i,n}{}^{\mathrm{T}},...,\mathbf{0}^{\mathrm{T}}]^{\mathrm{T}};\bar{\boldsymbol{s}}) & \text{if } \boldsymbol{s}_{i,n} \succeq \mathbf{0} \\ +\infty & \text{if } \boldsymbol{s}_{i,n} \not\succeq \mathbf{0} \end{cases}$$

$$F_{i,n}^*(\boldsymbol{y}) \triangleq \sup_{\boldsymbol{s}_{i,n}} \{\boldsymbol{y}^{\mathrm{T}} \boldsymbol{s}_{i,n} - F_{i,n}(\boldsymbol{s}_{i,n})\}, \quad (22)$$

where $\boldsymbol{s}_{i,n} \succeq \mathbf{0}$ denotes the element wise inequality, and where $\mathbf{0}$ is an all-zeros vector. Function $F_{i,n}^*$ is bounded from above if and only if all elements in $\boldsymbol{y}$ are strictly negative, such that its domain is given by $\operatorname{dom} F_{i,n}^* = \{\boldsymbol{y} \in \mathbb{R}^{N_i\times K} \mid \boldsymbol{y}_{k,n} < 0\}$. Moreover, for each $\boldsymbol{y} \in \operatorname{dom} F_{i,n}^*$, the solution set of the optimization problem defining $F_{i,n}^*$ is a singleton. Continuous differentiability of $F_{i,n}^*$ then immediately follows from [19, Theorem 11.8] and [19, Corollary 9.19]. As it is possible to equivalently define the dual function from (17) as

$$g_i(\boldsymbol{\lambda}^{\mathrm{p}}, \boldsymbol{\lambda}^{\mathrm{r}}) = \sum_{n\in\mathcal{N}_i} F_{i,n}^*(\boldsymbol{y}_n) + \boldsymbol{y}_0, \quad (23)$$

where each $\boldsymbol{y}_n$, $n \in \{0\} \cup \mathcal{N}_i$ is some affine vector function of the Lagrange multipliers $\boldsymbol{\lambda}^{\mathrm{p}}$ and $\boldsymbol{\lambda}^{\mathrm{r}}$, it follows that the dual function itself is continuously differentiable.

As the weights in the WSR problem are determined by the NUM algorithm that is chosen to update $\boldsymbol{\omega}$ in Algorithm 1, they may often be very different from one another. These large differences result in the curvature of the dual problem being strongly anisotropic, in the sense that the dual function's Hessian has a fairly high condition number. When applying a gradient descent method, this often leads to slow convergence in practice. It is therefore advised to solve problem (21) with a variable metric method such as L-BFGS-B [20].

V. Simulation results

In this section, simulation results are provided that illustrate the convergence behavior of NUM-DSB. Two versions of NUM-DSB are evaluated: one employing Jacobi updating, which for the original DSB often leads to faster convergence and solutions that achieve a higher WSR value, and one that employs GS updating, which enjoys stronger convergence guarantees.

The DSL systems under consideration contain 2, 3, or 4 vectoring groups. The DSLAM of each vectoring group connects to a maximum of 10 users, where the distance to the DSLAM ranges from $110\,\mathrm{m}$ for user 1 up to $200\,\mathrm{m}$ for user 10, increasing with $10\,\mathrm{m}$ for each consecutive user. The DSL systems with $N_i \leq 10$ contain, for each DSLAM, the first N_i users of their 10-user systems. The utility function of the minimal delay violation (MDV) scheduler [8] is considered, which is of the form $U(\boldsymbol{r}) = \sum_{i\in\mathcal{I}} \sum_{n\in\mathcal{N}_i} -a_{i,n}/r_{i,n}$ and represents an prediction of the number of delay violations as a function of the assigned rates. Note that $a_{i,n}$ is a constant that is determined by the cross-layer scheduler. Parameter settings for the DSL system are summarized in TABLE I.

The number of iterations required by NUM-DSB to converge is given in TABLE II, and the utility function value after convergence is displayed in Fig. 1. The algorithm is terminated when the relative decrease of the utility function between two iterations drops below 10^{-6}. From TABLE II, it is seen that the required number of iterations often decreases when more groups or users are added to the system. This can be explained by the increased interference in the system, rendering frequency division multiple access solutions optimal. In these cases, NUM-DSB on many tones switches off all but one vectoring group early on, and does not reactivate these in subsequent iterations, leading to fast convergence.

Overall, it is seen that GS updating leads to faster convergence than Jacobi updating. Moreover, Fig. 1 illustrates that GS updating also outperforms Jacobi updating in the sense that it achieves equal or higher utility function values. Contrary to the original DSB, it is no longer true that Jacobi updating mostly outperforms GS updating. In general, NUM-DSB requires only few iterations to obtain a local solution.

VI. Conclusion

In this paper, a fast algorithm has been presented for finding a local solution to the NUM problem, which is referred to

$$s^{\star}_{k,i,n} = \left[\frac{f_{\rm s}\omega_{i,n}/\log(2)}{\boldsymbol{t}_{k,i,n}\boldsymbol{\Lambda}^{\rm p}\boldsymbol{t}_{k,i,n} - f_{\rm s}\sum_{j\in\mathcal{I}\setminus\{i\}}\sum_{m\in\mathcal{N}_j}(\omega_{j,m}+\lambda^{\rm r}_{j,m})\frac{\partial b_{k,j,m}}{\partial s_{k,i,n}}\big|_{\boldsymbol{x}=\bar{\boldsymbol{x}}}} - \Gamma\left(\sigma_{k,i,n} + \sum_{j\in\mathcal{I}\setminus\{i\}}\sum_{m\in\mathcal{N}_j}|\hat{h}_{k,ij,nm}|^2\bar{s}_{k,j,m}\right)\right]^+ \quad (18)$$

TABLE I
G.FAST PARAMETER SETTINGS

Parameter	Value	Parameter	Value
$P^{n,\text{tot}}$	4 dBm	K	2047
f_s	48 kHz	Δ_f	51.75 kHz
Γ	10 dB	a^n	$10^8 \quad \forall n \in \mathcal{N}$

TABLE II
NUMBER OF ITERATIONS OF NUM-DSB BEFORE CONVERGENCE

	Gauss-Seidel			**Jacobi**		
N_i	$I=2$	$I=3$	$I=4$	$I=2$	$I=3$	$I=4$
1	53	5	8	59	8	7
2	49	22	38	52	6	14
3	19	29	29	53	18	17
4	17	36	29	52	19	20
5	19	24	9	55	20	21
6	23	9	10	53	21	15
7	22	39	7	60	20	15
8	23	36	8	62	7	16
9	31	47	9	91	8	23
10	49	46	10	70	9	22

as NUM-DSB, and which consists of iteratively solving sets of surrogate problems. The algorithm has been shown to be applicable in many DSL deployment scenarios, and regardless of the utility function's properties. In the context of grouped vectoring deployment scenarios, NUM-DSB leads to surrogate problems that are straightforward to solve. Moreover, using empirical evidence it has been shown that NUM-DSB needs only few iterations to find a satisfactory solution.

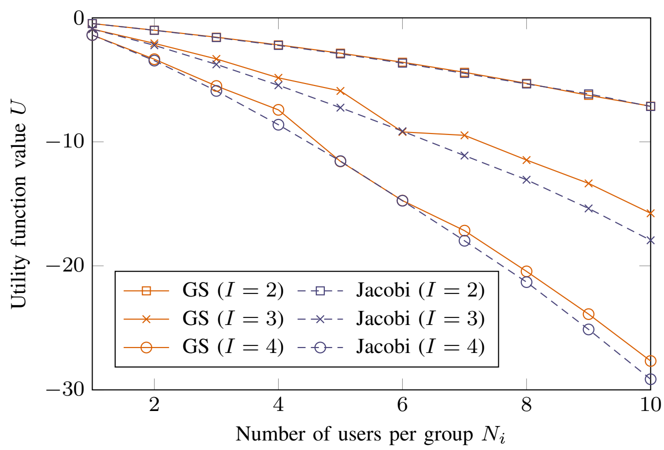

Fig. 1. Utility function value after convergence of NUM-DSB.

REFERENCES

[1] "Very high speed digital subscriber line transceivers 2 (VDSL2)," Tech. Rep. ITU-T G.993.2, 2007.

[2] R. Strobel, "Copper transmission for multi-gigabit hybrid copper-fiber access networks," in *Proc. of the Int. Workshop on Fiber Opt. in Access Networks (FOAN)*, Munich, Germany, Nov. 2017, pp. 1–6.

[3] W. Lanneer, J. Verdyck, P. Tsiaflakis, J. Maes, and M. Moonen, "Vectoring-based dynamic spectrum management for G.fast multi-user full-duplex transmission," in *Proc. of the Annu. Int. Symp. on Personal, Indoor, and Mobile Radio Commun. (PIMRC)*, Montreal, QC, Canada, Oct. 2017, pp. 1–5.

[4] M. Wolkerstorfer and C. F. Mecklenbrauker, "Variational inequality approach to spectrum balancing in vectoring xDSL networks," in *Proc. of the IEEE Int. Conf. on Commun. (ICC)*, May 2017, pp. 1–7.

[5] J. Brehmer, *Utility Maximization in Nonconvex Wireless Systems*, ser. Springer Found. in Signal Process., Commun. and Networking, 2012.

[6] H. Mehmood, M. Udell, and J. Cioffi, "Revenue Maximization for Broadband Service Providers Using Revenue Capacity," in *Proc. of the IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–7.

[7] Guocong Song, Ye Li, and L. Cimini, "Joint channel- and queue-aware scheduling for multiuser diversity in wireless OFDMA networks," *IEEE Trans. on Commun.*, vol. 57, no. 7, pp. 2109–2121, Jul. 2009.

[8] J. Van den Eynde, J. Verdyck, M. Moonen, and C. Blondia, "A delay-based cross-layer scheduler for adaptive DSL," in *Proc. of the IEEE Int. Conf. on Commun. (ICC)*, May 2017, pp. 1–6.

[9] P. Tsiaflakis, Y. Yi, M. Chiang, and M. Moonen, "Throughput and Delay Performance of DSL Broadband Access with Cross-Layer Dynamic Spectrum Management," *IEEE Trans. on Commun.*, vol. 60, no. 9, pp. 2700–2711, Sep. 2012.

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

[11] Zhi-Quan Luo and Shuzhong Zhang, "Dynamic Spectrum Management: Complexity and Duality," *IEEE J. of Select. Topics in Signal Process.*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[12] P. Tsiaflakis, M. Diehl, and M. Moonen, "Distributed Spectrum Management Algorithms for Multiuser DSL Networks," *IEEE. Trans. on Signal Process.*, vol. 56, no. 10, pp. 4825–4843, Oct. 2008.

[13] P. Tsiaflakis, R. B. Moraes, and M. Moonen, "A low-complexity algorithm for joint spectrum and signal coordination in upstream DSL transmission," in *Proc. of the IEEE Symp. on the Commun. and Veh. Technol. in the Benelux (SCVT)*, Nov. 2011, pp. 1–6.

[14] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-Minimization Algorithms in Signal Processing, Communications, and Machine Learning," *IEEE Trans. on Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.

[15] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A Unified Convergence Analysis of Block Successive Minimization Methods for Nonsmooth Optimization," *SIAM J. on Optimization*, vol. 23, no. 2, pp. 1126–1153, Jan. 2013.

[16] G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J. S. Pang, "Decomposition by Partial Linearization: Parallel Optimization of Multi-Agent Systems," *IEEE Trans. on Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.

[17] F. Facchinei, G. Scutari, and S. Sagratella, "Parallel Selective Algorithms for Nonconvex Big Data Optimization," *IEEE Trans. on Signal Process.*, vol. 63, no. 7, pp. 1874–1889, Apr. 2015.

[18] P. Tsiaflakis, F. Glineur, and M. Moonen, "Real-Time Dynamic Spectrum Management for Multi-User Multi-Carrier Communication Systems," *IEEE Trans. on Commun.*, vol. 62, no. 3, pp. 1124–1137, Mar. 2014.

[19] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*. Springer Science & Business Media, 2009, vol. 317.

[20] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A Limited Memory Algorithm for Bound Constrained Optimization," *SIAM J. on Scientific Computing*, vol. 16, no. 5, pp. 1190–1208, Sep. 1995.