KU LEUVEN

STADIUS
Center for Dynamical Systems, Signal Processing and Data Analytics

Citation/Reference	J. Verdyck, W. Lanneer, P. Tsiaflakis, W. Coomans, P. Patrinos, and M, Moonen (2019), **Optimal Dynamic Spectrum Management Algorithms for Multi-User Full-Duplex DSL** IEEE Access, vol. 7, Aug. 2019, 106600-106616
Archived version	As published
Published version	http://dx.doi.org/10.1109/ACCESS.2019.2926616
Journal homepage	https://ieeeaccess.ieee.org/
Author contact	jeroen.verdyck@esat.kuleuven.be *+ 32 (0)16 324723*
Abstract	Driven by the exceedingly high data rates achieved in single-user implementations, interest in a multi-user (MU) full-duplex (FDX) transmission for digital subscriber line (DSL) networks is surging. However, near-end crosstalk (NEXT) is no longer avoided in such networks, and hence, appropriate dynamic spectrum management (DSM) techniques are needed. Therefore, this paper proposes three novel DSM algorithms for the MU FDX DSL network. First, an optimal spectrum balancing (OSB) algorithm is derived that calculates the globally optimal resource allocation but does so at an exceedingly high computational cost. The key to this algorithm is a novel multiple access channel broadcast channel (MAC-BC) duality for the specific case of perfect NEXT cancellation at the distribution point unit. The two low-complexity distributed spectrum balancing (DSB) algorithms are then proposed, for which simulations show that their performance is very close to what is

	achieved by the OSB algorithm. Therefore, these DSB algorithms can be used to estimate the achievable performance of an MU FDX DSL network. Such performance estimations show that the FDX transmission can indeed lead to significant gains in MU DSL networks as well.
IR	Klik hier als u tekst wilt invoeren.

(article begins on next page)

IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 15, 2019, accepted June 25, 2019, date of publication July 8, 2019, date of current version August 16, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2926616

Optimal Dynamic Spectrum Management Algorithms for Multi-User Full-Duplex DSL

JEROEN VERDYCK[1], (Student Member, IEEE), WOUTER LANNEER[1], (Member, IEEE), PASCHALIS TSIAFLAKIS[2], (Member, IEEE), WERNER COOMANS[2], (Member, IEEE), PANAGIOTIS PATRINOS[1], (Member, IEEE), AND MARC MOONEN[1], (Fellow, IEEE)

[1]STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, Department of Electrical Engineering (ESAT), KU Leuven, 3000 Leuven, Belgium

[2]Copper Access and Indoor Team, Nokia Bell Labs, 2018 Antwerp, Belgium

Corresponding author: Jeroen Verdyck (jeroen.verdyck@esat.kuleuven.be)

This research work was carried out at the ESAT Laboratory of KU Leuven, in the frame of Fonds de la Recherche Scientifique (FNRS) and Fonds Wetenschappelijk Onderzoek Vlaanderen (FWO) EOS Project nr. 30452698 "(MUSE-WINET) MUlti-SErvice WIreless NETwork", FWO Research Project nr. G.0B1818N "Real-time adaptive cross-layer dynamic spectrum management for fifth generation broadband copper access networks", Vlaams Agentschap Innoveren & Ondernemen (VLAIO) O&O Project nr. HBC.2016.0055 "5GBB Fifth generation broadband access", VLAIO O&O Project nr. HBC.2017.1007 "(MIA) Multi-gigabit Innovations in Access".

ABSTRACT Driven by the exceedingly high data rates achieved in single-user implementations, interest in a multi-user (MU) full-duplex (FDX) transmission for digital subscriber line (DSL) networks is surging. However, near-end crosstalk (NEXT) is no longer avoided in such networks, and hence, appropriate dynamic spectrum management (DSM) techniques are needed. Therefore, this paper proposes three novel DSM algorithms for the MU FDX DSL network. First, an optimal spectrum balancing (OSB) algorithm is derived that calculates the globally optimal resource allocation but does so at an exceedingly high computational cost. The key to this algorithm is a novel multiple access channel broadcast channel (MAC-BC) duality for the specific case of perfect NEXT cancellation at the distribution point unit. The two low-complexity distributed spectrum balancing (DSB) algorithms are then proposed, for which simulations show that their performance is very close to what is achieved by the OSB algorithm. Therefore, these DSB algorithms can be used to estimate the achievable performance of an MU FDX DSL network. Such performance estimations show that the FDX transmission can indeed lead to significant gains in MU DSL networks as well.

INDEX TERMS DSL, dynamic spectrum management, G.mgfast, multi-user full-duplex, vectoring.

I. INTRODUCTION

With the potential of doubling the spectral efficiency, full-duplex (FDX) transmission for digital subscriber line (DSL) networks is receiving increased attention. In early stages of lab testing, single-user FDX has exceeded expectations. In XG-FAST trials at Nokia Bell Labs in Antwerp, for instance, aggregate data rates of 8.8 Gbit/s have been achieved on a single 30 m copper line [2]. When combined with multi-line bonding, FDX promises aggregate data rates even exceeding 20 Gbit/s. Consequently, FDX transmission will be part of G.mgfast, the upcoming ITU-T recommendation for DSL.

This paper studies FDX transmission in *multi-user* (MU) DSL networks, with a topology as illustrated in Fig. 1. The core network of the Internet service provider (ISP) is connected to the distribution point unit (DPU) through an optical fiber cable. In turn, the DPU is connected to N network terminations (NTs), one for each user, where each connection is established by means of a single twisted pair cable. At the DPU side, these twisted pair cables are bundled together in a cable binder. The dense packing of twisted pair cables in the cable binder results in an electromagnetic coupling, giving rise to interference or crosstalk. If not addressed appropriately, crosstalk can severely deteriorate the DSL network's performance.

A distinction is often made between near-end crosstalk (NEXT), i.e. interference generated by DPU (respectively NT) transmitters into neighboring DPU (NT) receivers, and far-end crosstalk (FEXT), i.e. interference generated by DPU (respectively NT) transmitters into NT (DPU) receivers at the other side of the DSL network. In previous DSL technologies, the effects of NEXT and FEXT have been managed in a dissimilar fashion.

The associate editor coordinating the review of this manuscript and approving it for publication was Yan Huo.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see http://creativecommons.org/licenses/by/4.0/

FEXT is typically dealt with by dynamic spectrum management (DSM) techniques. These techniques are often classified into two categories: spectrum coordination and signal coordination. Spectrum coordination involves jointly managing the transmit powers of different users, and signal coordination or vectoring comprises coordinating multiple users on a signal level [3]. Signal coordination techniques require the modems of different users to be co-located, thus introducing a difference between vectoring for upstream (US) transmission, i.e. NT to DPU, and for downstream (DS) transmission, i.e. DPU to NT.

In previous DSL technologies, the influence of NEXT has mostly[1] been avoided by dividing resources between US and DS transmission using FDD [4]–[7] or TDD [8]. This paper however considers FDX transmission, in which US and DS transmissions occur at the same time and in the same frequency band. In this FDX setting, NEXT is no longer avoided and should be dealt with using DSM techniques. At the DPU, interfering DS transmitters and victim US receivers are co-located. Therefore, the effects of NEXT can be reduced with signal coordination techniques such as NEXT cancellation. For the DSL setting, the DPU-side NEXT is typically not much stronger than the received signal, making it reasonable to assume perfect NEXT cancellation at the DPU such that US receivers indeed experience no NEXT. At the NTs however, interfering US transmitters and victim DS receivers are *not* co-located, and only spectrum coordination techniques are available to mitigate the NEXT impact. Remarkably, the assumption of perfect DPU-side NEXT cancellation provides for a more favorable problem structure, which is exploited in this paper to derive an optimal DSM algorithm for MU FDX networks.

The considered MU FDX DSL networks thus use joint vectoring and spectrum coordination to deal with FEXT, spectrum coordination to deal with NT-side NEXT, and echo cancellation to deal with DPU-side NEXT. The DSM algorithms developed in this paper determine optimal US and DS transmit powers for all users, as well as optimal receive filter and precoding vectors. Similar DSM algorithms have been considered in [9], [10]. In [9], DSM algorithms have been derived for FDX DSL networks with imperfect DPU-side NEXT cancellation, yielding a problem statement that is more general than what is considered in this paper. The derived algorithms are however not able to find the globally optimal DSM strategy. In [10] perfect DPU-side NEXT cancellation has been considered, but the derived zero-forcing precoders and receive filters are suboptimal. Lastly, in [11] a practical frame structure is proposed for MU FDX DSL networks by means of alternating between two FDX operating points, which can be provided by any DSM algorithm.

DSM algorithms for MU FDX networks have hitherto mostly been studied in the context of wireless systems. Relevant approaches include the WMMSE-based methods from [12]–[14], the successive convex approximation (SCA)-based methods from [15]–[17], and methods relying on a simplifying precoding/receive filter design [18]–[20]. In wireless systems however, the self-interference [2] (SI) power is orders of magnitude stronger than the received signal power, such that it cannot be assumed that all SI is canceled. The emphasis in [13]–[20] is therefore strongly on residual SI modeling and mitigation.

MAIN CONTRIBUTIONS

This paper introduces three novel DSM algorithms for MU FDX DSL networks. A first algorithm extends previously developed optimal spectrum balancing (OSB) algorithms for US and DS transmission [21]–[23] to an OSB algorithm for FDX transmission. Key is that the developed MAC-BC duality theory [3] yields a novel FDX-OSB algorithm, which is able to compute the globally optimal resource allocation. This optimality result is in contrast with the OSB-type algorithm developed in [9], which does not exploit perfect DPU-side NEXT cancellation and, as such, cannot guarantee global optimality.

The resulting FDX-OSB algorithm, however, exhibits an exponential complexity in the number of users, such that it becomes impractical for larger DSL networks. To overcome this problem, algorithms with a polynomial complexity in the number of users are presented. Both algorithms are based on previously developed distributed spectrum balancing (DSB) algorithms [24], [25]. The first algorithm, referred to as FDX-DD-DSB, is obtained by replacing the most demanding operation of FDX-OSB by an inexact low-complexity procedure. As FDX-DD-DSB is very similar to FDX-OSB, the performance of the two algorithms is anticipated to be comparable as well. The second algorithm, referred to as FDX-PD-DSB, is an SCA-based method that is not founded on the newly developed MAC-BC duality theory. As such, FDX-PD-DSB does not require perfect DPU-side NEXT cancellation, and can be applied the resource allocation problems from [13]–[17] as well. FDX-PD-DSB yields sub-problems that can be solved efficiently without relying on a solver as in [15]–[17]. Simulations demonstrate that FDX-PD-DSB exhibits better convergence characteristics than the WMMSE algorithms as in [13], [14]. In simulations, it is demonstrated that the low-complexity FDX-DSB algorithms achieve a similar performance as FDX-OSB at only a fraction of its computational cost. Lastly, simulations demonstrate that FDX as such can indeed lead to significant performance gains in MU DSL networks.

NOTATION

Upper case and lower case bold face symbols respectively denote matrices and vectors. The $N \times 1$ vector of zeros (respectively ones) and the identity matrix are denoted as

[1] ADSL and ADSL2 did support FDX transmission, but referred to it as 'echo canceling mode' [4], [5].

[2] In wireless networks, DPU-side NEXT is referred to as self-interference.

[3] MAC and BC respectively abbreviate *multiple access channel* and *broadcast channel*.

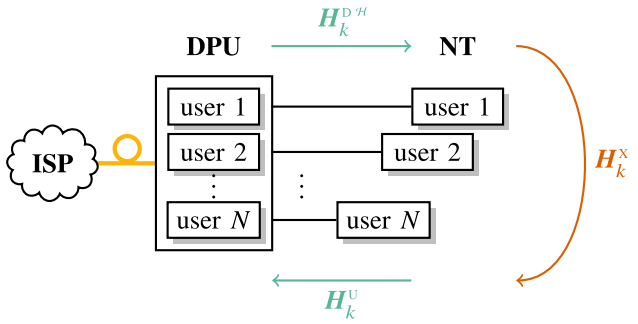

FIGURE 1. DSL network topology. The ISP core network is connected to the DPU through an optical fiber cable. In turn, the DPU is connected to *N* NTs, where each connection is established by means of a single twisted pair cable.

$\mathbf{0}_N$ ($\mathbf{1}_N$) and $\mathbf{I}$. Furthermore, $\boldsymbol{e}_n$ is the n-th vector in the standard basis of $\mathbb{R}^N$. The transpose, Hermitian transpose, complex conjugate, and inverse of the Hermitian transpose of a matrix $\mathbf{A}$ are respectively denoted as $\mathbf{A}^{\mathcal{T}}$, $\mathbf{A}^{\mathcal{H}}$, $\mathbf{A}^{*}$, and $\mathbf{A}^{-\mathcal{H}}$. The trace operator is denoted as $\mathrm{tr}(\cdot)$, and $\mathrm{diag}(\cdot)$ returns the diagonal matrix with the non-zero elements given by its argument. The null space of a matrix is denoted as $\mathrm{null}(\cdot)$. The Hadamard product of matrices $\mathbf{A}$ and $\mathbf{B}$ is denoted as $\mathbf{A} \circ \mathbf{B}$. The 2-norm and Frobenius norm are denoted as $\|\cdot\|_2$ and $\|\cdot\|_F$. Finally, the expected value operator is denoted as $IE[\cdot]$, and both the cardinality of a set and the modulus of a complex number are denoted as $|\cdot|$.

II. SYSTEM MODEL

DSL employs discrete multi-tone modulation (DMT) to split the available spectrum into a set of K orthogonal sub-carriers, often referred to as tones in DSL literature. It is assumed that no inter-carrier interference (ICI) is present, such that transmission can be modeled on each tone independently. Per tone channel models are presented for US and DS transmission in Sections II-A and II-B. Performance metrics for the DSL network are presented in Section II-C.

A. UPSTREAM CHANNEL

US transmission in an N-user FDX DSL network with no ICI and with perfect DPU-side NEXT cancellation can be modeled on each tone as a multiple access channel (MAC), i.e.

$$\boldsymbol{y}_k^{\mathrm{U}} = \mathbf{H}_k^{\mathrm{U}} \boldsymbol{x}_k^{\mathrm{U}} + \boldsymbol{z}_k^{\mathrm{U}} \quad \forall k. \tag{1}$$

In the channel equation, the superscript 'U' indicates that a US variable is considered. Furthermore, the vector $\boldsymbol{x}_k^{\mathrm{U}} = \left[x_{k,1}^{\mathrm{U}}, \ldots, x_{k,N}^{\mathrm{U}}\right]^{\mathcal{T}}$ contains the transmitted signals of all the users on tone k. Vectors $\boldsymbol{z}_k^{\mathrm{U}}$ and $\boldsymbol{y}_k^{\mathrm{U}}$ denote the additive Gaussian noise and received signal, and have the same size as $\boldsymbol{x}_k^{\mathrm{U}}$. Moreover, $\mathbf{H}_k^{\mathrm{U}}$ denotes the $N \times N$ US channel matrix, with $\left[\mathbf{H}_k^{\mathrm{U}}\right]_{nm} = h_{k,nm}^{\mathrm{U}}$ the transfer function between transmitter m and receiver n evaluated at tone k.

The average symbol power of user n on tone k is defined as $s_{k,n}^{\mathrm{U}} = \Delta_{\mathrm{f}} IE\left[|x_{k,n}^{\mathrm{U}}|^2\right]$, with Δ_{f} the tone spacing. For each user n, the total US transmit power is then given by

$$P_n^{\mathrm{U}} = \sum_k s_{k,n}^{\mathrm{U}}. \tag{2}$$

In the above equation, $\sum_k$ is a shorthand notation for the summation over all elements $k \in \{1, \ldots, K\}$. Similarly, $\sum_n$ and $\prod_n$ respectively denote the summation and (Cartesian) product over all elements $n \in \mathcal{N} = \{1, \ldots, N\}$, and $\sum_{m \neq n}$ denotes the summation over all elements $m \in \mathcal{N} \setminus \{n\}$. Furthermore, $\boldsymbol{\Sigma}_k^{\mathrm{U}} = \Delta_{\mathrm{f}} IE\left[\boldsymbol{z}_k^{\mathrm{U}} \boldsymbol{z}_k^{\mathrm{U}\,\mathcal{H}}\right]$ denotes the US additive noise covariance matrix.

In US transmission, signal coordination is possible at the DPU. The non-linear general decision feedback equalizer (GDFE) receiver structure is assumed, i.e. transmitted symbols are iteratively estimated from the received signal as

$$\hat{x}_{k,n}^{\mathrm{U}} = \boldsymbol{r}_{k,n}^{\mathrm{U}\,\mathcal{H}} \left(\boldsymbol{y}_k^{\mathrm{U}} - \sum_{m<n} \boldsymbol{h}_{k,m}^{\mathrm{U}} \hat{x}_{k,m}^{\mathrm{U}} \right). \tag{3}$$

In the above equation, it is assumed that the decoding order is determined beforehand and w.l.o.g. given by the user index order. Moreover, $\mathbf{R}_k^{\mathrm{U}}$ is the tone k receive filter matrix with $\boldsymbol{r}_{k,n}^{\mathrm{U}}$ its n-th column. Likewise, $\boldsymbol{h}_{k,n}^{\mathrm{U}}$ denotes the n-th column of $\mathbf{H}_k^{\mathrm{U}}$. The resulting signal-to-interference-plus-noise ratio (SINR) for user n on tone k is given by

$$\gamma_{k,n}^{\mathrm{U}}(\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{r}_{k,n}^{\mathrm{U}}, \boldsymbol{\Sigma}_k^{\mathrm{U}}) = \frac{s_{k,n}^{\mathrm{U}} |\boldsymbol{r}_{k,n}^{\mathrm{U}\,\mathcal{H}} \boldsymbol{h}_{k,n}^{\mathrm{U}}|^2}{\sum_{m>n} s_{k,m}^{\mathrm{U}} |\boldsymbol{r}_{k,n}^{\mathrm{U}\,\mathcal{H}} \boldsymbol{h}_{k,m}^{\mathrm{U}}|^2 + \boldsymbol{r}_{k,n}^{\mathrm{U}\,\mathcal{H}} \boldsymbol{\Sigma}_k^{\mathrm{U}} \boldsymbol{r}_{k,n}^{\mathrm{U}}} \tag{4}$$

with $\boldsymbol{s}_k^{\mathrm{U}} = [s_{k,1}^{\mathrm{U}}, \ldots, s_{k,N}^{\mathrm{U}}]^{\mathcal{T}}$. Dependencies will always be explicitly mentioned for the SINR, as well as for bit loading variables later on. Lastly, it is noted that when a linear receiver structure is considered, the feedback term is to be removed from (3) and the summation in the denominator of (4) should be over $\mathcal{N} \setminus \{n\}$ instead of $m > n$.

In (4), it is seen that only the SINR of user n depends on $\boldsymbol{r}_{k,n}^{\mathrm{U}}$. The minimum mean square error (MMSE) receiver is therefore optimal, as it maximizes the SINR of user n, which is given by (5) where $\boldsymbol{\Psi}_{k,n}^{\mathrm{U}} = \sum_{m>n} s_{k,m}^{\mathrm{U}} \boldsymbol{h}_{k,m}^{\mathrm{U}} \boldsymbol{h}_{k,m}^{\mathrm{U}\,\mathcal{H}} + \boldsymbol{\Sigma}_k^{\mathrm{U}}$ is user n's received interference-plus-noise covariance matrix on tone k. Substituting (5) into (4), $\mathbf{R}_k^{\mathrm{U}}$ can be eliminated from the US SINR expression, yielding (6). In case a linear receiver structure is considered, the summation in the definition of $\boldsymbol{\Psi}_{k,n}^{\mathrm{U}}$ should be over $\mathcal{N} \setminus \{n\}$ instead of $m > n$, and then (5) and (6) still apply.

$$\boldsymbol{r}_{k,n}^{\mathrm{U}} = \left(\boldsymbol{\Psi}_{k,n}^{\mathrm{U}}\right)^{-1} \boldsymbol{h}_{k,n}^{\mathrm{U}} \tag{5}$$

$$\gamma_{k,n}^{\mathrm{U}}(\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{\Sigma}_k^{\mathrm{U}}) = s_{k,n}^{\mathrm{U}} \boldsymbol{h}_{k,n}^{\mathrm{U}\,\mathcal{H}} \left(\boldsymbol{\Psi}_{k,n}^{\mathrm{U}}\right)^{-1} \boldsymbol{h}_{k,n}^{\mathrm{U}} \tag{6}$$

B. DOWNSTREAM CHANNEL

DS transmission in an N-user FDX DSL network with no ICI can be modeled on each tone independently as a broadcast channel (BC) with an additional term for the NEXT caused by the NTs.

$$\boldsymbol{y}_k^{\mathrm{D}} = \mathbf{H}_k^{\mathrm{D}\,\mathcal{H}} \tilde{\boldsymbol{x}}_k^{\mathrm{D}} + \mathbf{H}_k^{\mathrm{X}\,\mathcal{H}} \boldsymbol{x}_k^{\mathrm{U}} + \boldsymbol{z}_k^{\mathrm{D}} \tag{7}$$

In the channel equation, the superscript 'D' indicates that DS variables are considered. Vectors $\tilde{\boldsymbol{x}}_k^{\mathrm{D}}, \boldsymbol{y}_k^{\mathrm{D}}$, and $\boldsymbol{z}_k^{\mathrm{D}}$ have the same size as $\boldsymbol{x}_k^{\mathrm{U}}$ and denote the transmitted signal, received signal, and additive Gaussian noise. Moreover, $\mathbf{H}_k^{\mathrm{D}}$ and $\mathbf{H}_k^{\mathrm{X}}$ denote the Hermitian transpose of the $N \times N$ DS channel matrix and the $N \times N$ NT-side NEXT channel matrix, with $\left[\mathbf{H}_k^{\mathrm{D}}\right]_{mn} = {h_{k,nm}^{\mathrm{D}}}^*$ and $\left[\mathbf{H}_k^{\mathrm{X}}\right]_{mn} = {h_{k,nm}^{\mathrm{X}}}^*$ the complex conjugate of the transfer function between transmitter m and receiver n evaluated on tone k. The Hermitian transpose of the channel matrices are used in (7) to simplify notation later on.

In DS transmission, signal coordination is possible at the transmitter. Transmitted signals are generated using (8) where $\mathbf{T}_k^{\mathrm{D}}$ is the precoding matrix. Furthermore, the average symbol power of user n on tone k is defined as $s_{k,n}^{\mathrm{D}} = \Delta_{\mathrm{f}} IE\left[|x_{k,n}^{\mathrm{D}}|^2\right]$. For each line, the total DS transmit power is then given by (9) where $\mathbf{S}_k^{\mathrm{D}} = \mathrm{diag}(\boldsymbol{s}_k^{\mathrm{D}})$ with $\boldsymbol{s}_k^{\mathrm{D}} = [s_{k,1}^{\mathrm{D}}, \ldots, \boldsymbol{s}_{k,N}^{\mathrm{D}}]^{\mathcal{T}}$. The received noise power for user n on tone k is $\sigma_{k,n}^{\mathrm{D}} = \Delta_{\mathrm{f}} IE\left[|z_{k,n}^{\mathrm{D}}|^2\right]$.

$$\tilde{\boldsymbol{x}}_k^{\mathrm{D}} = \mathbf{T}_k^{\mathrm{D}} \boldsymbol{x}_k^{\mathrm{D}} \tag{8}$$

$$P_n^{\mathrm{D}} = \sum_k \left[\mathbf{T}_k^{\mathrm{D}} \mathbf{S}_k^{\mathrm{D}} \mathbf{T}_k^{\mathrm{D}\,\mathcal{H}}\right]_{nn} \tag{9}$$

A non-linear transmitter structure implementing dirty paper coding (DPC) is assumed [26], such as the Tomlinson-Harashima precoder. DPC-based transmitters successively encode the symbols of different users, treating interference generated by previously encoded users as side information and treating the interference generated by other users as noise [27]. According to the original DPC results by Costa [28], considering the interference generated by previously encoded users to be side information allows one to obtain the same performance as when this interference were not present. The resulting SINR of user n on tone k is given as

$$\gamma_{k,n}^{\mathrm{D}}(\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{D}}, \mathbf{T}_k^{\mathrm{D}}, \boldsymbol{\sigma}_k^{\mathrm{D}}) = \frac{s_{k,n}^{\mathrm{D}} |\boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}} \boldsymbol{t}_{k,n}^{\mathrm{D}}|^2}{\sum_{m<n} s_{k,m}^{\mathrm{D}} |\boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}} \boldsymbol{t}_{k,m}^{\mathrm{D}}|^2 + \boldsymbol{h}_{k,n}^{\mathrm{XX}\,\mathcal{T}} \boldsymbol{s}_k^{\mathrm{U}} + \sigma_{k,n}^{\mathrm{D}}} \tag{10}$$

where $\boldsymbol{h}_{k,n}^{\mathrm{XX}}$ is the n-th column of the power transfer matrix of the NEXT channel on tone k, which is defined as $[\mathbf{H}_k^{\mathrm{XX}}]_{nm} = |h_{k,nm}^{\mathrm{X}}|^2$. In equation (10), it has been assumed that the user encoding order is fixed beforehand and w.l.o.g. given by the reversed user index order. When a linear transmitter structure is considered, all interference terms are treated as noise and the summation in the denominator of (10) should be over $\mathcal{N} \setminus \{n\}$ instead of $m < n$.

C. PERFORMANCE METRICS

When the number of users N in a DSL network is large, the interference-plus-noise received by each user is well approximated by a Gaussian distribution. Under this assumption, the relation between the SINR γ and the achievable bit loading b, which is assumed to be a continuous variable, is accurately modeled by (11) where $\log_2(\cdot)$ is the binary logarithm and where the SNR-gap to capacity Γ accounts for the difference in performance between ideal Gaussian signaling and the practical modulation and coding scheme in use. The Γ is additionally determined by the employed noise margin and by the target bit error rate. Typical values for Γ lie between 9.5 dB and 10.5 dB. The data rate of user n is calculated using (12) where f_{s} is the symbol rate.

$$b(\gamma) = \log_2\left(1 + \Gamma^{-1}\gamma\right) \tag{11}$$

$$R_n = f_{\mathrm{s}} \sum_k b(\gamma_{k,n}) \tag{12}$$

III. FULL-DUPLEX DYNAMIC SPECTRUM MANAGEMENT

The so-called rate-adaptive DSM problem is considered, which optimizes the network's performance by selecting the resource allocation $\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{D}}, \mathbf{T}_k^{\mathrm{D}}$ for all tones k that solves the weighted sum rate (WSR) maximization problem, i.e.

$$\underset{\substack{\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{D}} \in \mathbb{R}_+^N \,\forall k \\ \mathbf{T}_k^{\mathrm{D}} \,\forall k}}{\text{maximize}} \sum_n \omega_n^{\mathrm{U}} \sum_k b(\gamma_{k,n}^{\mathrm{U}}) + \sum_n \omega_n^{\mathrm{D}} \sum_k b(\gamma_{k,n}^{\mathrm{D}}) \tag{13a}$$

$$\text{subject to } P_n^{\mathrm{U}} \leq P_{\mathrm{tot}}^{\mathrm{U}},\ \forall n \text{ and } P_n^{\mathrm{D}} \leq P_{\mathrm{tot}}^{\mathrm{D}},\ \forall n. \tag{13b}$$

In (13a), the factor f_{s} has been omitted for brevity, and $\mathbb{R}_+^N \triangleq \{\boldsymbol{s} \in \mathbb{R}^N \mid s_n \geq 0,\ \forall n\}$. By adjusting the real positive weights ω in the objective function (13a), all Pareto-optimal resource allocations can be obtained [21].

The considered WSR maximization problem (13) is subject to US an DS total per line power constraints (13b), and to positivity constraints. Spectral mask and bit cap constraints are not explicitly accounted for in the formulation of problem (13). Although it is possible to add these constraints to the WSR maximization problem, their absence allows for a notation that is not too unwieldy in Section IV and Section V. Yet another version of problem (13) can be obtained by abandoning the assumption of perfect DPU-side NEXT cancellation. After the derivation of each algorithm, it will be indicated which modifications have to be made to adapt the algorithms to these modified problem statements.

IV. OPTIMAL SPECTRUM BALANCING

In this section, the FDX-OSB algorithm is developed, which finds the globally optimal DSM strategy for an MU FDX DSL network. Contrary to the OSB-type algorithm developed in [9], the favorable structure of the problem considered here enables the development of a novel MAC-BC duality theory that, in turn, allows finding the globally optimal resource allocation. The FDX-OSB algorithm contains the MAC-OSB algorithm [22] and BC-OSB algorithm [23] as special cases, and relies on dual decomposition, MAC-BC duality, and an exhaustive grid search. Each of these components is now addressed individually.

A. DUAL DECOMPOSITION

The idea of dual decomposition is to solve the Lagrange dual problem associated with problem (13), as defined in (14).

Problem (14) is often referred to as the master problem, and minimizes the Lagrange dual function $q(\boldsymbol{\lambda}^{\mathrm{U}}, \boldsymbol{\lambda}^{\mathrm{D}})$ with respect to the US and DS Lagrange multipliers $\boldsymbol{\lambda}^{\mathrm{U}}$ and $\boldsymbol{\lambda}^{\mathrm{D}}$. In turn, the Lagrange dual function is defined as the maximum of the Lagrangian over all possible resource allocations, as defined in (15), which is referred to as the slave problem.

$$\underset{\boldsymbol{\lambda}^{\mathrm{U}}, \boldsymbol{\lambda}^{\mathrm{D}} \in \mathbb{R}_{+}^{N}}{\text{minimize}} \quad q(\boldsymbol{\lambda}^{\mathrm{U}}, \boldsymbol{\lambda}^{\mathrm{D}}) \tag{14}$$

$$q(\boldsymbol{\lambda}^{\mathrm{U}}, \boldsymbol{\lambda}^{\mathrm{D}}) = \max_{\substack{\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{D}} \in \mathbb{R}_{+}^{N} \, \forall k \\ \mathbf{T}_k^{\mathrm{D}} \, \forall k}} \left\{ \mathcal{L} \right\} \tag{15}$$

It is noted that in (15), the resource allocations are no longer subject to total per line power constraints. These constraints are incorporated into the Lagrangian, which is defined as

$$\mathcal{L} = \sum_k \left(\mathcal{L}_k^{\mathrm{U}} + \mathcal{L}_k^{\mathrm{D}} \right) + \sum_n \left(\lambda_n^{\mathrm{U}} P_{\mathrm{tot}}^{\mathrm{U}} + \lambda_n^{\mathrm{D}} P_{\mathrm{tot}}^{\mathrm{D}} \right) \tag{16a}$$

$$\mathcal{L}_k^{\mathrm{U}} = \sum_n \omega_n^{\mathrm{U}} b\left(\gamma_{k,n}^{\mathrm{U}}(\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{\Sigma}_k^{\mathrm{U}}) \right) - \boldsymbol{\lambda}^{\mathrm{U}\,\mathcal{T}} \boldsymbol{s}_k^{\mathrm{U}} \tag{16b}$$

$$\mathcal{L}_k^{\mathrm{D}} = \sum_n \omega_n^{\mathrm{D}} b\left(\gamma_{k,n}^{\mathrm{D}}(\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{D}}, \mathbf{T}_{k,n}^{\mathrm{D}}, \boldsymbol{\sigma}_k^{\mathrm{D}}) \right) - \sum_n s_{k,n}^{\mathrm{D}} \boldsymbol{t}_{k,n}^{\mathrm{D}\,\mathcal{H}} \boldsymbol{\Lambda}^{\mathrm{D}} \boldsymbol{t}_{k,n}^{\mathrm{D}}, \tag{16c}$$

with $\boldsymbol{\Lambda}^{\mathrm{D}} = \mathrm{diag}(\boldsymbol{\lambda}^{\mathrm{D}})$. Assuming the number of tones K is large, the time sharing property of [29] holds such that the duality gap [30] between problem (13) and problem (15) is zero. The master problem in (14) is convex but non-smooth, and can therefore be solved using a subgradient-based scheme such as the subgradient method or the ellipsoid method. Subgradients are calculated as

$$\begin{aligned} \boldsymbol{g}^{\mathrm{U}} &= \boldsymbol{P}^{\mathrm{U}} - \mathbf{1}_N P_{\mathrm{tot}}^{\mathrm{U}}, \\ \boldsymbol{g}^{\mathrm{D}} &= \boldsymbol{P}^{\mathrm{D}} - \mathbf{1}_N P_{\mathrm{tot}}^{\mathrm{D}}. \end{aligned} \tag{17}$$

For each new value of the Lagrange multipliers $\boldsymbol{\lambda}^{\mathrm{U}}$ and $\boldsymbol{\lambda}^{\mathrm{D}}$, the slave problem (15) is solved to obtain $\boldsymbol{P}^{\mathrm{U}}$ and $\boldsymbol{P}^{\mathrm{D}}$. Each term $\mathcal{L}_k^{\mathrm{U}} + \mathcal{L}_k^{\mathrm{D}}$ in (15) depends on tone k resource allocation variables only. The slave problem is therefore separable across tones, such that the Lagrange dual function can be evaluated by solving K decoupled per tone slave problems of the following form.

$$\underset{\substack{\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{D}} \in \mathbb{R}_{+}^{N} \\ \mathbf{T}_k^{\mathrm{D}}}}{\text{maximize}} \quad \mathcal{L}_k^{\mathrm{U}} + \mathcal{L}_k^{\mathrm{D}} \tag{18}$$

Section IV-B and Section IV-C elaborate on how the global optimum of each per tone slave problem can be obtained.

B. MAC-BC DUALITY FOR FDX DSL

Whereas $\mathbf{R}_k^{\mathrm{U}}$ is effectively removed from the optimization problem due to the favorable structure of $\gamma_{k,n}^{\mathrm{U}}$ in (6), the same does not go for $\mathbf{T}_k^{\mathrm{D}}$. Therefore, a new MAC-BC duality theory for FDX networks is now developed, which allows transforming $\mathcal{L}_k^{\mathrm{D}}$ into an equivalent dual US Lagrangian plus a coupling bilinear term. The proposed theory is based on [23], [31], [32]. Using the duality theory, $\mathbf{T}_k^{\mathrm{D}}$ can then be eliminated from $\gamma_{k,n}^{\mathrm{D}}$ in the same way as $\mathbf{R}_k^{\mathrm{U}}$ is removed from $\gamma_{k,n}^{\mathrm{U}}$.

The dual US Lagrangian is defined in (19), with the expression for the dual upstream SINR given by (20).

$$\mathcal{L}_k^{\mathrm{dU}} = \sum_n \omega_n^{\mathrm{D}} b\left(\gamma_{k,n}^{\mathrm{dU}}(\boldsymbol{s}_k^{\mathrm{dU}}, \boldsymbol{t}_{k,n}^{\mathrm{D}}, \boldsymbol{\Lambda}^{\mathrm{D}}) \right) - \boldsymbol{\sigma}_k^{\mathrm{D}\,\mathcal{T}} \boldsymbol{s}_k^{\mathrm{dU}} \tag{19}$$

$$\gamma_{k,n}^{\mathrm{dU}}(\boldsymbol{s}_k^{\mathrm{dU}}, \boldsymbol{t}_{k,n}^{\mathrm{D}}, \boldsymbol{\Lambda}^{\mathrm{D}}) = \frac{s_{k,n}^{\mathrm{dU}} |\boldsymbol{t}_{k,n}^{\mathrm{D}\,\mathcal{H}} \boldsymbol{h}_{k,n}^{\mathrm{D}}|^2}{\sum_{m>n} s_{k,m}^{\mathrm{dU}} |\boldsymbol{t}_{k,n}^{\mathrm{D}\,\mathcal{H}} \boldsymbol{h}_{k,m}^{\mathrm{D}}|^2 + \boldsymbol{t}_{k,n}^{\mathrm{D}\,\mathcal{H}} \boldsymbol{\Lambda}^{\mathrm{D}} \boldsymbol{t}_{k,n}^{\mathrm{D}}} \tag{20}$$

The superscript dU indicates that a variable of the dual upstream channel is considered. By comparing (20) to (4), and (19) to (16b), it is seen that this dual US Lagrangian corresponds to a MAC system with channel matrix $\mathbf{H}_k^{\mathrm{D}}$, receive filter matrix $\mathbf{T}_k^{\mathrm{D}}$, noise covariance matrix $\boldsymbol{\Lambda}^{\mathrm{D}}$, Lagrange multiplier values $\boldsymbol{\sigma}_k^{\mathrm{D}}$, and a decoding order that is the reversed encoding order of the original DS system. Duality between $\mathcal{L}_k^{\mathrm{dU}}$ and $\mathcal{L}_k^{\mathrm{D}}$ is established in the following proposition.

Proposition 1 (MAC-BC duality for FDX DSL): Let $\boldsymbol{s}_k^{\mathrm{U}} \in \mathbb{R}_{+}^{N}$ and $\mathbf{T}_k^{\mathrm{D}}$ such that $\boldsymbol{t}_{k,n}^{\mathrm{D}} \neq \mathbf{0}_N \; \forall n$. Furthermore, assume that both $\boldsymbol{\lambda}^{\mathrm{D}}$ and $\boldsymbol{\sigma}_k^{\mathrm{D}}$ are strictly positive. The following statements then hold true.

- **P.1** *For each dual symbol power vector $\boldsymbol{s}_k^{dU} \in \mathbb{R}_{+}^{N}$, a symbol power vector $\boldsymbol{s}_k^{D} \in \mathbb{R}_{+}^{N}$ exists such that equalities* (21) *and* (22) *are satisfied.*
- **P.2** *The converse is also true, i.e. for each symbol power vector $\boldsymbol{s}_k^{D} \in \mathbb{R}_{+}^{N}$, a dual symbol power vector $\boldsymbol{s}_k^{dU} \in \mathbb{R}_{+}^{N}$ exists such that equalities* (21) *and* (22) *are satisfied.*

$$\gamma_{k,n}^{\mathrm{D}}(\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{D}}, \mathbf{T}_k^{\mathrm{D}}, \boldsymbol{\sigma}_k^{\mathrm{D}}) = \gamma_{k,n}^{\mathrm{dU}}(\boldsymbol{s}_k^{\mathrm{dU}}, \boldsymbol{t}_{k,n}^{\mathrm{D}}, \boldsymbol{\Lambda}^{\mathrm{D}}), \; \forall n \tag{21}$$

$$\mathcal{L}_k^{\mathrm{D}} = \mathcal{L}_k^{\mathrm{dU}} - \boldsymbol{s}_k^{\mathrm{U}\,\mathcal{T}} \mathbf{H}_k^{\mathrm{XX}} \boldsymbol{s}_k^{\mathrm{dU}} \tag{22}$$

Proof: Following the reasoning in [32], the validity of **P.1** is confirmed by construction of $\boldsymbol{s}_k^{\mathrm{D}}$ from $\boldsymbol{s}_k^{\mathrm{dU}}$. **P.2** can then be proven analogously.

The symbol power vector $\boldsymbol{s}_k^{\mathrm{D}}$ corresponding to a given $\boldsymbol{s}_k^{\mathrm{dU}}$ is found by solving the following system of equations, which is constructed from and equivalent to the equalities in (21).

$$\mathbf{Z}_k^{\mathrm{dU}} \boldsymbol{s}_k^{\mathrm{D}} = \left(\mathbf{H}_k^{\mathrm{XX}\,\mathcal{T}} \boldsymbol{s}_k^{\mathrm{U}} + \boldsymbol{\sigma}_k^{\mathrm{D}} \right) \circ \boldsymbol{s}_k^{\mathrm{dU}} \tag{23}$$

$$\left[\mathbf{Z}_k^{\mathrm{dU}} \right]_{nm} = \begin{cases} -\boldsymbol{t}_{k,m}^{\mathrm{D}\,\mathcal{H}} \left(s_{k,n}^{\mathrm{dU}} \boldsymbol{h}_{k,n}^{\mathrm{D}} \boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}} \right) \boldsymbol{t}_{k,m}^{\mathrm{D}} & \text{if } n < m \\ \boldsymbol{t}_{k,m}^{\mathrm{D}\,\mathcal{H}} \boldsymbol{\Psi}_{k,m}^{\mathrm{dU}} \boldsymbol{t}_{k,m}^{\mathrm{D}} & \text{if } n = m \\ 0 & \text{if } n > m \end{cases}$$

In (23), $\boldsymbol{\Psi}_{k,n}^{\mathrm{dU}} = \sum_{m>n} s_{k,m}^{\mathrm{dU}} \boldsymbol{h}_{k,m}^{\mathrm{U}} \boldsymbol{h}_{k,m}^{\mathrm{D}\,\mathcal{H}} + \boldsymbol{\Lambda}^{\mathrm{D}}$ is the dual interference-plus-noise covariance matrix of user n. Positivity of $\boldsymbol{s}_k^{\mathrm{D}}$ follows from $\mathbf{Z}_k^{\mathrm{dU}}$ being an M-matrix, such that its inverse contains only non-negative elements [32], [33]. Furthermore, summing the rows of (23) yields

$$\sum_n s_{k,n}^{\mathrm{D}} \boldsymbol{t}_{k,n}^{\mathrm{D}\,\mathcal{H}} \boldsymbol{\Lambda}^{\mathrm{D}} \boldsymbol{t}_{k,n}^{\mathrm{D}} = \boldsymbol{\sigma}_k^{\mathrm{D}\,\mathcal{T}} \boldsymbol{s}_k^{\mathrm{dU}} + \boldsymbol{s}_k^{\mathrm{U}\,\mathcal{T}} \mathbf{H}_k^{\mathrm{XX}} \boldsymbol{s}_k^{\mathrm{dU}}, \tag{24}$$

confirming that the equality in (22) holds. □

In order to adapt this duality result to a linear receiver structure, the summation in the denominator of (20), as well

as the summation in the definition of $\mathbf{\Psi}_{k,n}^{\mathrm{dU}}$, should be over $\mathcal{N} \setminus \{n\}$ instead of $m < n$. Furthermore, the lower triangular elements of $\mathbf{Z}_k^{\mathrm{dU}}$ should be defined similarly to the upper triangular elements.

Corollary 2: **P.2** *also holds when* $\boldsymbol{t}_{k,n}^{\mathrm{D}} = \mathbf{0}_N$ *for any* n.

Proof: If $\boldsymbol{t}_{k,n}^{\mathrm{D}} = \mathbf{0}_N$, then an equivalent DS resource allocation can be obtained that achieves the same values for $\mathcal{L}_k^{\mathrm{D}}$ and $\gamma_{k,n}^{\mathrm{D}}(\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{D}}, \mathbf{T}_k^{\mathrm{D}}, \boldsymbol{\sigma}_k^{\mathrm{D}})$ $\forall n$, by setting $\boldsymbol{t}_{k,n}^{\mathrm{D}} \neq \mathbf{0}_N$ and $s_{k,n}^{\mathrm{D}} = 0$. This equivalent resource allocation satisfies the assumptions of Proposition 1. □

Corollary 3 (Dual MAC optimality): The solution to

$$\underset{\boldsymbol{s}_k^{U}, \boldsymbol{s}_k^{dU} \in \mathbb{R}_+^N, \mathbf{T}_k^{D}}{\text{maximize}} \quad \mathcal{L}_k^{U} + \mathcal{L}_k^{dU} - \boldsymbol{s}_k^{U\,\mathcal{T}} \mathbf{H}_k^{XX} \boldsymbol{s}_k^{dU}, \tag{25}$$

when transformed to the DS domain using (23)*, is the solution to the original per tone slave problem in* (18).

Proof: The Corollary is proven by contradiction. Assume w.l.o.g. that a solution to problem (25) exists with $\boldsymbol{t}_{k,n}^{\mathrm{D}} \neq \mathbf{0}_N$ $\forall n$. If Corollary 3 were false, a solution to problem (18) would exist achieving a higher value for $\mathcal{L}_k^{\mathrm{U}} + \mathcal{L}_k^{\mathrm{D}}$ than the transformed solution of (25). Corollary 2 would then imply that a dual MAC domain resource allocation exists achieving a value for $\mathcal{L}_k^{\mathrm{U}} + \mathcal{L}_k^{\mathrm{dU}} - \boldsymbol{s}_k^{\mathrm{U}\,\mathcal{T}} \mathbf{H}_k^{\mathrm{XX}} \boldsymbol{s}_k^{\mathrm{dU}}$ that is higher than the value achieved by the solution to (25), which is a contradiction. □

In problem (25), $\mathbf{T}_k^{\mathrm{D}}$ can now be eliminated from $\gamma_{k,n}^{\mathrm{dU}}$ in the same way as $\mathbf{R}_k^{\mathrm{U}}$ has been eliminated from $\gamma_{k,n}^{\mathrm{U}}$ in (4). The optimal dual domain receive filter vector $\boldsymbol{t}_{k,n}^{\mathrm{D}}$ is given by (26) and substituting this expression into (20) yields the expression for the SINR in (27).

$$\boldsymbol{t}_{k,n}^{\mathrm{D}} = \left(\mathbf{\Psi}_{k,n}^{\mathrm{dU}}\right)^{-1} \boldsymbol{h}_{k,n}^{\mathrm{D}} \tag{26}$$

$$\gamma_{k,n}^{\mathrm{dU}}(\boldsymbol{s}_k^{\mathrm{dU}}, \mathbf{\Lambda}^{\mathrm{D}}) = s_{k,n}^{\mathrm{dU}} \boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}} \left(\mathbf{\Psi}_{k,n}^{\mathrm{dU}}\right)^{-1} \boldsymbol{h}_{k,n}^{\mathrm{D}} \tag{27}$$

C. EXHAUSTIVE GRID SEARCH

After applying MAC-BC duality and eliminating both $\boldsymbol{R}_k^{\mathrm{U}}$ and $\boldsymbol{T}_k^{\mathrm{D}}$, the following simplified per tone slave problem is obtained.

$$\underset{\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{dU}} \in \mathbb{R}_+^N}{\text{maximize}} \quad \mathcal{L}_k^{\mathrm{U}} + \mathcal{L}_k^{\mathrm{dU}} - \boldsymbol{s}_k^{\mathrm{U}\,\mathcal{T}} \mathbf{H}_k^{\mathrm{XX}} \boldsymbol{s}_k^{\mathrm{dU}} \tag{28}$$

In problem (28), $\mathcal{L}_k^{\mathrm{U}}$ depends only on $\boldsymbol{s}_k^{\mathrm{U}}$, whereas $\mathcal{L}_k^{\mathrm{dU}}$ depends only on $\boldsymbol{s}_k^{\mathrm{dU}}$. This non-convex problem is solved by performing an exhaustive search over a grid of possible bit loading combinations $\prod_n \mathcal{B} \times \prod_n \mathcal{B}$, with $\mathcal{B}$ the set of possible bit loading values (assumed to be the same in US and DS). As the number of points in the grid is $|\mathcal{B}|^{2N}$, the complexity of this exhaustive search is exponential in $2N$.

For each vector $[\boldsymbol{b}_k^{\mathrm{U}\,\mathcal{T}}\, \boldsymbol{b}_k^{\mathrm{D}\,\mathcal{T}}]^{\mathcal{T}}$ in the search grid, the corresponding US and DS symbol powers $\boldsymbol{s}_k^{\mathrm{U}}$ and $\boldsymbol{s}_k^{\mathrm{dU}}$ are calculated, and then the objective function of (28) is evaluated. It is noted that $\boldsymbol{b}_k^{\mathrm{U}}$ (respectively $\boldsymbol{b}_k^{\mathrm{D}}$) solely depends on $\boldsymbol{s}_k^{\mathrm{U}}$ ($\boldsymbol{s}_k^{\mathrm{dU}}$), not on $\boldsymbol{s}_k^{\mathrm{dU}}$ ($\boldsymbol{s}_k^{\mathrm{U}}$), such that $\boldsymbol{s}_k^{\mathrm{U}}$ and $\mathcal{L}_k^{\mathrm{U}}$ ($\boldsymbol{s}_k^{\mathrm{dU}}$ and $\mathcal{L}_k^{\mathrm{dU}}$) can be calculated from $\boldsymbol{b}_k^{\mathrm{U}}$ ($\boldsymbol{b}_k^{\mathrm{dU}}$) independently of $\boldsymbol{b}_k^{\mathrm{dU}}$ ($\boldsymbol{b}_k^{\mathrm{U}}$). By calculating $\boldsymbol{s}_k^{\mathrm{U}}$ and $\mathcal{L}_k^{\mathrm{U}}$ (respectively $\mathbf{H}_k^{\mathrm{XX}} \boldsymbol{s}_k^{\mathrm{dU}}$ and $\mathcal{L}_k^{\mathrm{dU}}$) for each $\boldsymbol{b} \in \prod_n \mathcal{B}$ beforehand, storing the obtained values in a set $\mathcal{A}^{\mathrm{U}}$ ($\mathcal{A}^{\mathrm{D}}$), the complexity of evaluating the objective function of (28) for each $[\boldsymbol{b}_k^{\mathrm{U}\,\mathcal{T}}\, \boldsymbol{b}_k^{\mathrm{D}\,\mathcal{T}}]^{\mathcal{T}} \in \prod_n \mathcal{B} \times \prod_n \mathcal{B}$ is kept to a minimum.

The symbol power vector $\boldsymbol{s}_k^{\mathrm{U}}$ is obtained from $\boldsymbol{b}_k^{\mathrm{U}}$ by first calculating each $\gamma_{k,n}^{\mathrm{U}}$ using (11), and by subsequently inverting (27). When $s_{k,m}^{\mathrm{U}}$ is known for all users $m > n$, $s_{k,n}^{\mathrm{U}}$ can be calculated in closed form as

$$s_{k,n}^{\mathrm{U}} = \frac{\gamma_{k,n}}{\boldsymbol{h}_{k,n}^{\mathrm{U}\,\mathcal{H}} (\mathbf{\Psi}_{k,n}^{\mathrm{U}})^{-1} \boldsymbol{h}_{k,n}^{\mathrm{U}}}. \tag{29}$$

Therefore, $s_{k,n}^{\mathrm{U}}$ is evaluated sequentially starting from user N and ending with user 1. The dual symbol power vector $\boldsymbol{s}_k^{\mathrm{dU}}$ can similarly be obtained from $\boldsymbol{b}_k^{\mathrm{D}}$. In case a linear transceiver structure is considered, $\boldsymbol{s}_k^{\mathrm{U}}$ ($\boldsymbol{s}_k^{\mathrm{dU}}$) can be obtained from $\boldsymbol{b}_k^{\mathrm{U}}$ ($\boldsymbol{b}_k^{\mathrm{D}}$) using an iterative calculation, as in [34].

The resulting algorithm is summarized in Algorithm 1, and is referred to as full-duplex optimal spectrum balancing (FDX-OSB). To the best of the authors' knowledge, FDX-OSB is the first algorithm that is guaranteed to find the globally optimal solution of a MU FDX resource allocation problem as in (13). The complexity of FDX-OSB mainly stems from line 14, requiring N multiplications and $N + 1$ additions in each iteration. The overall complexity of Algorithm 1 is therefore $\mathcal{O}(I_{\mathrm{SG}} K |\mathcal{B}|^{2N} N)$, with I_{SG} the number of iterations required by the subgradient algorithm to converge.

In the above derivation of FDX-OSB, it has been indicated how to adapt the formulas used in Algorithm 1 when a linear transceiver structure is considered. While Algorithm 1 applies to DSM problem (13), other versions of FDX-OSB are obtained by including bit cap and/or spectral mask constraints. Bit cap constraints are easily enforced by appropriately designing the bit loading set $\mathcal{B}$. US spectral

Algorithm 1 FDX-OSB

1: Choose $\mu > 0$, $\boldsymbol{\lambda}^{\mathrm{U}} \in \mathbb{R}_+^N$, $\boldsymbol{\lambda}^{\mathrm{D}} \in \mathbb{R}_+^N$, $\mathcal{B} \subset \mathbb{N}$, and $\epsilon > 0$
2: **while** ($|P_n^{\mathrm{U}} - P_{\mathrm{tot}}^{\mathrm{U}}| > \delta_{\mathrm{P}}$ & $\lambda_n^{\mathrm{U}} > \delta_\lambda$ for any n) or ($|P_n^{\mathrm{D}} - P_{\mathrm{tot}}^{\mathrm{D}}| > \delta_{\mathrm{P}}$ & $\lambda_n^{\mathrm{D}} > \delta_\lambda$ for any n) **do**
3: **for** $k \in \mathcal{K}$ **do**
4: $\left\{\boldsymbol{s}_k^{\mathrm{U}}, \mathbf{R}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{D}}, \mathbf{T}_k^{\mathrm{D}}\right\} \leftarrow$ ExhSearch($\boldsymbol{\lambda}^{\mathrm{U}}, \boldsymbol{\lambda}^{\mathrm{D}}, k$)
5: Calculate P_n^{U}, P_n^{D} $\forall n$ using (2), (9)
6: Update $\boldsymbol{\lambda}^{\mathrm{U}}$, $\boldsymbol{\lambda}^{\mathrm{D}}$ using (17)
7: **function** ExhSearch($\boldsymbol{\lambda}^{\mathrm{U}}, \boldsymbol{\lambda}^{\mathrm{D}}, k$)
8: Initialize $\mathcal{A}^{\mathrm{U}} \leftarrow \emptyset$, $\mathcal{A}^{\mathrm{dU}} \leftarrow \emptyset$, $\mathcal{L}_k{}^{\star} \leftarrow -\inf$
9: **for all** $\boldsymbol{b} \in \prod_n \mathcal{B}$ **do**
10: Calculate $\boldsymbol{s}_k^{\mathrm{U}}$, $\boldsymbol{s}_k^{\mathrm{dU}}$ from $\boldsymbol{b}$ by inverting (6), (27)
11: Calculate $\mathcal{L}_k^{\mathrm{U}}$, $\mathcal{L}_k^{\mathrm{dU}}$ using (16b), (19)
12: Add $\left(\boldsymbol{s}_k^{\mathrm{U}}, \mathcal{L}_k^{\mathrm{U}}\right)$ to $\mathcal{A}^{\mathrm{U}}$, $\left(\mathbf{H}_k^{\mathrm{XX}} \boldsymbol{s}_k^{\mathrm{dU}}, \mathcal{L}_k^{\mathrm{dU}}\right)$ to $\mathcal{A}^{\mathrm{dU}}$
13: **for all** $(\boldsymbol{s}_k^{\mathrm{U}}, \mathcal{L}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{dU}}, \mathcal{L}_k^{\mathrm{dU}}) \in \mathcal{A}^{\mathrm{U}} \times \mathcal{A}^{\mathrm{dU}}$ **do**
14: Set $\mathcal{L}_k \leftarrow \mathcal{L}_k^{\mathrm{U}} + \mathcal{L}_k^{\mathrm{dU}} - \boldsymbol{s}_k^{\mathrm{U}\,\mathcal{T}} \mathbf{H}_k^{\mathrm{XX}} \boldsymbol{s}_k^{\mathrm{dU}}$
15: **if** $\mathcal{L}_k{}^{\star} < \mathcal{L}_k$ **then**
16: Set $\mathcal{L}_k{}^{\star} \leftarrow \mathcal{L}_k$, $\boldsymbol{s}_k^{\mathrm{U}\star} \leftarrow \boldsymbol{s}_k^{\mathrm{U}}$, $\boldsymbol{s}_k^{\mathrm{dU}\star} \leftarrow \boldsymbol{s}_k^{\mathrm{dU}}$
17: Calculate $\boldsymbol{R}_k^{\mathrm{U}\star}$, $\boldsymbol{T}_k^{\mathrm{D}\star}$ from $\boldsymbol{s}_k^{\mathrm{U}\star}$, $\boldsymbol{s}_k^{\mathrm{dU}\star}$ using (5), (26)
18: Calculate BC user powers $\boldsymbol{s}_k^{\mathrm{D}\star}$ by solving (23)

mask constraints are easily enforced by discarding $(\boldsymbol{s}_k^{\mathrm{U}}, \mathcal{L}_k^{\mathrm{U}})$-tuples that do not satisfy these constraints in line 12 of Algorithm 1. DS spectral mask constraints however are tougher to enforce. A possible method has been proposed in [35], consisting of dualizing the DS spectral mask constraints when formulating the Lagrange dual problem (14). Finally, problem (13) can be altered by relaxing the assumption of perfect DPU-side NEXT cancellation. In this case however, the solution strategy employed by FDX-OSB cannot be adopted to find the global optimum, as the proposed MAC-BC duality transformation is not able to eliminate $\boldsymbol{s}_k^{\mathrm{D}}$ from the resulting US Lagrangian $\mathcal{L}_k^{\mathrm{U}}$, and hence one should resort to suboptimal algorithms as in [9].

V. DISTRIBUTED SPECTRUM BALANCING

With FDX-OSB, finding the global optimum of problem (13) comes at the price of an exceedingly high complexity. In this section, two algorithms are presented that exhibit a lower complexity, but are only able to find a local optimum of problem (13). The algorithms are FDX adaptations of the distributed spectrum balancing (DSB) algorithm that has previously been developed for both the spectrum coordination problem [24] and for the US and DS joint signal and spectrum coordination problem [25], [34], [35]. A brief introduction to the DSB algorithm structure is given first, and then the two DSB algorithms are presented for the FDX channel.

A. DSB ALGORITHM

The DSB algorithm [24], [25] comprises an iterative procedure to solve optimization problems of the following form.

$$\begin{aligned} &\underset{\mathbf{X}}{\text{maximize}} \sum_{i=1}^{i_{\max}} f_i(\mathbf{X}) \\ &\text{subject to } \mathbf{X}_i \in \mathcal{X}_i,\ \forall i \end{aligned} \tag{30}$$

In problem (30), the decision variable $\mathbf{X}$ is comprised of multiple coordinate blocks $\mathbf{X}_i$, i.e. $\mathbf{X} = \left[\mathbf{X}_1{}^{\mathcal{T}}, \ldots, \mathbf{X}_{i_{\max}}^{\mathcal{T}}\right]^{\mathcal{T}}$. Each function f_i has a (possibly empty) coordinate block $\mathbf{X}_i$ associated with it in which f_i is concave. Moreover, f_i is convex in each $\mathbf{X}_j$ with $j \in \{1, \ldots, i_{\max}\} \setminus \{i\}$. In each iteration of the DSB algorithm the objective function of (30) is replaced by the sum of surrogate functions $\tilde{f}_i$, each depending only on a single coordinate block $\mathbf{X}_i$. The resulting surrogate problem has the following form.

$$\begin{aligned} &\underset{\mathbf{X}}{\text{maximize}} \sum_{i=1}^{i_{\max}} \tilde{f}_i(\mathbf{X}_i; \bar{\mathbf{X}}) \\ &\text{subject to } \mathbf{X}_i \in \mathcal{X}_i, \quad \forall i \end{aligned} \tag{31}$$

In (31), $\bar{\mathbf{X}}$ denotes the current value of $\mathbf{X}$. The surrogate functions $\tilde{f}_i$ are defined as

$$\tilde{f}_i(\mathbf{X}_i; \bar{\mathbf{X}}) \triangleq f_i\left([\bar{\mathbf{X}}_1{}^{\mathcal{T}}, \ldots, \mathbf{X}_i{}^{\mathcal{T}}, \ldots, \bar{\mathbf{X}}_{i_{\max}}{}^{\mathcal{T}}]^{\mathcal{T}}\right) - \mathrm{tr}(\mathbf{A}_i{}^{\mathcal{H}}\mathbf{X}_i) \tag{32}$$

with $\mathbf{A}_i = -\nabla_{\mathbf{X}_i{}^*}\left(\sum_{j\neq i} f_j(\mathbf{X})\right)\big|_{\mathbf{X}=\bar{\mathbf{X}}}$ the negative conjugate cogradient of $\sum_{j\neq i} f_j(\mathbf{X})$, and where $\mathbf{A}_i{}^{\mathcal{H}}\mathbf{X}_i$ is assumed to be Hermitian such that its trace is real. By iteratively solving surrogate problems (31), which are separable across coordinate blocks $\mathbf{X}_i$, the DSB algorithm can find a stationary point of the original problem as in (30).

It is noted that the acronym 'DSB' is adopted from [24], [25], but that the same algorithmic structure is also known as successive convex approximation (SCA). SCA algorithms that have DSB as a special case and for which theoretical convergence results are available include BSUM, SJBR, and FLEXA. Convergence results have first been established for the 'block successive upper-bound minimization' (BSUM) algorithm [36], which updates only a single coordinate block $\mathbf{X}_i$ after solving (31). This sequential mode of operation is referred to as Gauss-Seidel updating. More recently, convergence results have been established for Algorithm 1 from [37], hereafter referred to as SJBR, which updates all coordinate blocks by taking a step in the direction of the solution to (31). The parallel mode of operation employed by SJBR is referred to as Jacobi updating. It is noted that convergence of SJBR requires each $\tilde{f}_i$ to be strongly convex, guaranteeing that all subproblems as in (31) have a unique solution. Convergence has also been established for the 'inexact flexible parallel algorithm' (FLEXA) [38], which is similar to SJBR but can also be applied to non-smooth problems.[4]

B. FDX DUAL DOMAIN DSB ALGORITHM

The first FDX DSB algorithm is obtained by applying the DSB algorithm to problem (28). Three different types of functions f_i are identified.

$$f_n^{\mathrm{U}}(\boldsymbol{s}_k^{\mathrm{U}}) = \omega_n^{\mathrm{U}} b\left(\gamma_{k,n}^{\mathrm{U}}(\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{\Sigma}_k^{\mathrm{U}})\right) - \lambda_n^{\mathrm{U}} s_{k,n}^{\mathrm{U}} \tag{33a}$$

$$f_n^{\mathrm{dU}}(\boldsymbol{s}_k^{\mathrm{dU}}) = \omega_n^{\mathrm{D}} b\left(\gamma_{k,n}^{\mathrm{dU}}(\boldsymbol{s}_k^{\mathrm{dU}}, \boldsymbol{\Lambda}^{\mathrm{D}})\right) - \sigma_{k,n}^{\mathrm{D}} s_{k,n}^{\mathrm{dU}} \tag{33b}$$

$$f^{\mathrm{X}}(\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{s}_k^{\mathrm{dU}}) = -\boldsymbol{s}_k^{\mathrm{U}\,\mathcal{T}} \mathbf{H}_k^{\mathrm{XX}} \boldsymbol{s}_k^{\mathrm{dU}} \tag{33c}$$

The coordinate block associated with f_n^{U} is $s_{k,n}^{\mathrm{U}}$ and the coordinate block associated with f_n^{dU} is $s_{k,n}^{\mathrm{dU}}$. An empty coordinate block is associated with f^{X}. Functions f_n^{U}, f_n^{dU}, and f^{X} are concave in their associated coordinate blocks and convex in the other coordinate blocks. The gradients $\mathbf{A}_i$ in the corresponding surrogate functions are real scalars and calculated as

$$a_{k,n}^{\mathrm{U}} = \sum_{m<n} \boldsymbol{h}_{k,n}^{\mathrm{U}}{}^{\mathcal{H}} \bar{\boldsymbol{\Phi}}_{k,m}^{\mathrm{U}} \boldsymbol{h}_{k,n}^{\mathrm{U}} + [\mathbf{H}_k^{\mathrm{XX}}]_{\mathrm{row}\ m} \bar{\boldsymbol{s}}_k^{\mathrm{dU}} \tag{34a}$$

$$a_{k,n}^{\mathrm{dU}} = \sum_{m<n} \boldsymbol{h}_{k,n}^{\mathrm{D}}{}^{\mathcal{H}} \bar{\boldsymbol{\Phi}}_{k,m}^{\mathrm{dU}} \boldsymbol{h}_{k,n}^{\mathrm{D}} + \boldsymbol{h}_{k,n}^{\mathrm{XX}\,\mathcal{T}} \bar{\boldsymbol{s}}_k^{\mathrm{U}}, \tag{34b}$$

[4] The results in [36]–[38] are derived for problems with real variables. In [39] however, it is argued that same convergence results apply to real problems that are obtained from a complex problem by using separate variables for the real and imaginary parts of the complex variables. As is suggested in [39] however, one can equivalently work directly with complex variables by means of 'Wirtinger derivatives' (as will be done in Section V-C). For more information on optimization of real functions in complex variables, the reader is referred to [40].

Algorithm 2 Approximate Evaluation of (28) for FDX-DD-DSB

1: **function** DSBsearch($\boldsymbol{\lambda}^{\mathrm{U}}$, $\boldsymbol{\lambda}^{\mathrm{D}}$, k)
2: Choose $\bar{\boldsymbol{s}}_k^{\mathrm{U}} \in \mathbb{R}_+^N$, $\bar{\boldsymbol{s}}_k^{\mathrm{dU}} \in \mathbb{R}_+^N$, $\epsilon_\tau \in (0,1)$, set $\tau \leftarrow 1$
3: **while** $\mathcal{L}_k^{\mathrm{U}} + \mathcal{L}_k^{\mathrm{dU}} - \boldsymbol{s}_k^{\mathrm{U}\,\mathcal{T}} \mathbf{H}_k^{\mathrm{XX}} \boldsymbol{s}_k^{\mathrm{dU}}$ not converged **do**
4: Calculate $a_{k,n}^{\mathrm{U}}$ and $a_{k,n}^{\mathrm{dU}}$ $\forall n$ using (34)
5: Calculate $s_{k,n}^{\mathrm{U}}$ and $s_{k,n}^{\mathrm{dU}}$ $\forall n$ using (36)
6: Set $\bar{\boldsymbol{s}}_k^{\mathrm{U}} \leftarrow (1-\tau)\bar{\boldsymbol{s}}_k^{\mathrm{U}} + \tau \boldsymbol{s}_k^{\mathrm{U}}$
7: Set $\bar{\boldsymbol{s}}_k^{\mathrm{dU}} \leftarrow (1-\tau)\bar{\boldsymbol{s}}_k^{\mathrm{dU}} + \tau \boldsymbol{s}_k^{\mathrm{dU}}$
8: Set $\tau \leftarrow \tau(1-\tau\epsilon_\tau)$
9: Calculate $\boldsymbol{R}_k^{\mathrm{U}}$, $\boldsymbol{T}_k^{\mathrm{D}}$ from $\boldsymbol{s}_k^{\mathrm{U}}$, $\boldsymbol{s}_k^{\mathrm{dU}}$ using (5), (26)
10: Calculate $\boldsymbol{s}_k^{\mathrm{D}}$ by solving (23)

where $\boldsymbol{\Phi}_{k,m}^{\mathrm{U}}$ (respectively $\boldsymbol{\Phi}_{k,m}^{\mathrm{dU}}$) expresses the sensitivity of user m's weighted bit loading to changes in its (dual) interference-plus-noise covariance matrix $\boldsymbol{\Psi}_{k,m}^{\mathrm{U}}$ ($\boldsymbol{\Psi}_{k,m}^{\mathrm{dU}}$).

$$\boldsymbol{\Phi}_{k,n}^{\mathrm{U}} = \log(2)^{-1}\omega_n^{\mathrm{U}} \frac{s_{k,n}^{\mathrm{U}} (\boldsymbol{\Psi}_{k,n}^{\mathrm{U}})^{-1} \boldsymbol{h}_{k,n}^{\mathrm{U}} \boldsymbol{h}_{k,n}^{\mathrm{U}\,\mathcal{H}} (\boldsymbol{\Psi}_{k,n}^{\mathrm{U}})^{-1}}{\Gamma + s_{k,n}^{\mathrm{U}} \boldsymbol{h}_{k,n}^{\mathrm{U}\,\mathcal{H}} (\boldsymbol{\Psi}_{k,n}^{\mathrm{U}})^{-1} \boldsymbol{h}_{k,n}^{\mathrm{U}}} \tag{35a}$$

$$\boldsymbol{\Phi}_{k,n}^{\mathrm{dU}} = \log(2)^{-1}\omega_n^{\mathrm{D}} \frac{s_{k,n}^{\mathrm{dU}} (\boldsymbol{\Psi}_{k,n}^{\mathrm{dU}})^{-1} \boldsymbol{h}_{k,n}^{\mathrm{D}} \boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}} (\boldsymbol{\Psi}_{k,n}^{\mathrm{dU}})^{-1}}{\Gamma + s_{k,n}^{\mathrm{dU}} \boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}} (\boldsymbol{\Psi}_{k,n}^{\mathrm{dU}})^{-1} \boldsymbol{h}_{k,n}^{\mathrm{D}}} \tag{35b}$$

When a linear transceiver structure is considered, the summations in (34) should be over $\mathcal{N} \setminus \{n\}$ instead of $m < n$.

The resulting surrogate problem is separable across all variables, and can therefore be solved independently for each user. From the KKT conditions of the surrogate problem, a closed form solution for $s_{k,n}^{\mathrm{U}}$ and $s_{k,n}^{\mathrm{dU}}$ is obtained, given by

$$s_{k,n}^{\mathrm{U}} = \left[\frac{\log(2)^{-1}\omega_n^{\mathrm{U}}}{\lambda_n^{\mathrm{U}} + a_{k,n}^{\mathrm{U}}} - \frac{\Gamma}{\boldsymbol{h}_{k,m}^{\mathrm{U}\,\mathcal{H}} (\bar{\boldsymbol{\Psi}}_{k,m}^{\mathrm{U}})^{-1} \boldsymbol{h}_{k,m}^{\mathrm{U}}} \right]^+, \tag{36a}$$

$$s_{k,n}^{\mathrm{dU}} = \left[\frac{\log(2)^{-1}\omega_n^{\mathrm{D}}}{\sigma_{k,n}^{\mathrm{D}} + a_{k,n}^{\mathrm{dU}}} - \frac{\Gamma}{\boldsymbol{h}_{k,m}^{\mathrm{D}\,\mathcal{H}} (\bar{\boldsymbol{\Psi}}_{k,m}^{\mathrm{dU}})^{-1} \boldsymbol{h}_{k,m}^{\mathrm{D}}} \right]^+, \tag{36b}$$

where $[\cdot]^+ \triangleq \max\{\cdot, 0\}$.

The resulting algorithm, referred to as full-duplex dual domain distributed spectrum balancing (FDX-DD-DSB), is obtained from Algorithm 1 by replacing the exhaustive search function by Algorithm 2. The complexity of Algorithm 2 is determined by the calculation of $(\bar{\boldsymbol{\Psi}}_{k,n}^{\mathrm{U}})^{-1}$ and $(\bar{\boldsymbol{\Psi}}_{k,n}^{\mathrm{dU}})^{-1}$ in line 4, where the calculation of a single matrix inverse results in a complexity of $\mathcal{O}(N^3)$. However, by using the Sherman-Morrison identity [41], $(\bar{\boldsymbol{\Psi}}_{k,n}^{\mathrm{U}})^{-1}$ can be obtained from $(\bar{\boldsymbol{\Psi}}_{k,n+1}^{\mathrm{U}})^{-1}$ with a complexity of only $\mathcal{O}(N^2)$. All other calculations in lines 4-8 of Algorithm 2 have a per user complexity of at most $\mathcal{O}(N^2)$. Each iteration of Algorithm 2 therefore has a complexity of $\mathcal{O}(N^3)$. This results in a total complexity of $\mathcal{O}(I_{\mathrm{SG}} K I_{\mathrm{DSB}} N^3)$ for FDX-DD-DSB, where I_{DSB} is the number of iterations required by the DSB algorithm to converge.

Algorithm 2 adopts the SJBR step size rule for updating the symbol power variables in lines 6-8. As such, FDX-DD-DSB is a special case of the SJBR algorithm [37], such that it inherits SJBR's convergence properties. SJBR converges to a stationary point of problem (30), provided that the following convergence conditions are satisfied [37]:

- **C.1** each $\mathcal{X}_i$ is closed and convex;
- **C.2** each f_i is continuously differentiable;
- **C.3** the gradient of each f_i is Lipschitz continuous;
- **C.4** $\sum_{i=1}^{i_{\max}} f_i$ is coercive with respect to $\prod_i \mathcal{X}_i$;
- **C.5** each surrogate $\tilde{f}_i$ is uniformly strongly convex for all $\bar{\mathbf{X}}$.

Convergence conditions **C.1** and **C.2** are trivially satisfied. Moreover, **C.3** follows from $\boldsymbol{\Phi}_{k,n}^{\mathrm{U}}$ and $\boldsymbol{\Phi}_{k,n}^{\mathrm{dU}}$ being continuously differentiable. Assuming $\lambda_n^{\mathrm{U}} > 0$ for all n, **C.4** and **C.5** are satisfied as the set of possible solutions to (36a) and (36b) is upper bounded by $(\lambda_n^{\mathrm{U}} \log(2))^{-1}\omega_n^{\mathrm{U}}$ and $(\sigma_n^{\mathrm{D}} \log(2))^{-1}\omega_n^{\mathrm{D}}$ [37]. In case $\lambda_n^{\mathrm{U}} = 0$ for any n, **C.4**-**C.5** are possibly not satisfied, and one should resort to Gauss-Seidel iterations with $\epsilon_\tau = 0$. The resulting algorithm is then a special case of the BSUM algorithm, which has more lenient convergence conditions.

In [29], [34], it has been pointed out that the subgradient method, which FDX-DD-DSB uses to update $\boldsymbol{\lambda}^{\mathrm{U}}$ and $\boldsymbol{\lambda}^{\mathrm{D}}$, is not guaranteed to converge when $q(\boldsymbol{\lambda}^{\mathrm{U}}, \boldsymbol{\lambda}^{\mathrm{D}})$ is evaluated only approximately. As Algorithm 2 does not necessarily return the global maximizer of problem (28), FDX-DD-DSB is not guaranteed to converge. FDX-DD-DSB has however been observed to converge in every scenario for which it has been executed.

In the derivation of FDX-DD-DSB, it has already been indicated how to adapt Algorithm 2 when a linear transceiver structure is considered. Other versions of FDX-DD-DSB can be obtained by including bit cap or spectral mask constraints. Bit cap constraints are readily enforced by using the following augmented bit loading function $b_{\mathrm{cap}}(\gamma) = \max\{b(\gamma), b_{\max}\}$ or, equivalently, by upper bounding $s_{k,n}^{\mathrm{U}}$ (or $s_{k,n}^{\mathrm{dU}}$) as in [35]. Convergence of the Gauss-Seidel version of FDX-DD-DSB is still guaranteed in this case, be it not to a stationary point. US spectral mask constraints are easily enforced by replacing $[\cdot]^+$ in (36a) with $[\cdot]_0^{s_k^{\mathrm{mask}}} = \max\{\min\{\cdot, s_k^{\mathrm{mask}}\}, 0\}$. This constraint enforces **C.4** and **C.5**, guaranteeing that Algorithm 2 converges to a stationary point. DS spectral mask constraints, however, are tougher to enforce. A possible method has been proposed in [35], consisting of dualizing the DS spectral mask constraints when formulating the Lagrange dual problem. In this case however, the time sharing property cannot be invoked to warrant a zero duality gap leading, in some cases, to problematic convergence behavior. Finally, problem (13) can be altered by relaxing the assumption of perfect DPU-side NEXT cancellation. In this case however, as for FDX-OSB, the employed solution strategy cannot be adopted here, as the MAC-BC duality transfor-

mation is not able to eliminate s_k^{D} from the resulting US Lagrangian $\mathcal{L}_k^{\mathrm{U}}$.

C. FDX PRIMAL DOMAIN DSB ALGORITHM

A second FDX DSB algorithm is obtained by applying the DSB algorithm to a reformulated version of the original optimization problem (13), similar to [37], [42]. This reformulation is obtained by introducing DS transmit covariance matrices $\mathbf{Q}_{k,n}^{\mathrm{D}} = s_{k,n}^{\mathrm{D}} \boldsymbol{t}_{k,n}^{\mathrm{D}} \boldsymbol{t}_{k,n}^{\mathrm{D}\,\mathcal{H}}$. It is seen that any transmit covariance matrix must be positive semidefinite ($\mathbf{Q}_{k,n}^{\mathrm{D}} \succeq 0$), Hermitian ($\mathbf{Q}_{k,n}^{\mathrm{D}\,\mathcal{H}} = \mathbf{Q}_{k,n}^{\mathrm{D}}$), and rank one. Employing $\mathbf{Q}_{k,n}^{\mathrm{D}}$, the SINR for user n can be expressed as

$$\gamma_{k,n}^{\mathrm{D}}(\boldsymbol{s}_k^{\mathrm{U}}, \mathbf{Q}_k^{\mathrm{D}}, \sigma_{k,n}^{\mathrm{D}}) = (\psi_{k,n}^{\mathrm{D}})^{-1} \boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}} \mathbf{Q}_{k,n}^{\mathrm{D}} \boldsymbol{h}_{k,n}^{\mathrm{D}}, \tag{37}$$

with $\psi_{k,n}^{\mathrm{D}} = \sum_{m<n} \boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}} \mathbf{Q}_{k,m}^{\mathrm{D}} \boldsymbol{h}_{k,n}^{\mathrm{D}} + \boldsymbol{h}_{k,n}^{\mathrm{XX}\,\mathcal{H}} \boldsymbol{s}_k^{\mathrm{U}} + \sigma_{k,n}^{\mathrm{D}}$ the received interference-plus-noise power and with $\mathbf{Q}_k^{\mathrm{D}} = [\mathbf{Q}_{k,1}^{\mathrm{D}\,\mathcal{T}}, \ldots, \mathbf{Q}_{k,N}^{\mathrm{D}\,\mathcal{T}}]^{\mathcal{T}}$. The total DS transmit power for line n is given by

$$P_n^{\mathrm{D}} = \sum_k \sum_m [\mathbf{Q}_{k,m}^{\mathrm{D}}]_{nn}. \tag{38}$$

Using (37) and (38), $\boldsymbol{t}_{k,n}^{\mathrm{D}}$ and $s_{k,n}^{\mathrm{D}}$ can be fully eliminated from (13), which then becomes a function of $\mathbf{Q}_{k,n}^{\mathrm{D}}$ instead. It is noted that, even though the reformulated optimization problem will not explicitly assert that $\mathbf{Q}_{k,n}^{\mathrm{D}}$ should be rank one, any $\mathbf{Q}_{k,n}^{\mathrm{D}}$ yielded by FDX-PD-DSB does indeed satisfy this condition (see proof Lemma 4). When a linear transceiver structure is considered, the summation in the definition of $\psi_{k,n}^{\mathrm{D}}$ should be over $\mathcal{N} \setminus \{n\}$ instead of $m < n$.

In the reformulated optimization problem, two different types of functions f_i are identified.

$$f_n^{\mathrm{U}}(\boldsymbol{s}_1^{\mathrm{U}}, \ldots, \boldsymbol{s}_K^{\mathrm{U}}) = \omega_n^{\mathrm{U}} \sum_k b\big(\gamma_{k,n}^{\mathrm{U}}(\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{\Sigma}_k^{\mathrm{U}})\big) \tag{39a}$$

$$f_n^{\mathrm{D}}(\boldsymbol{s}_1^{\mathrm{U}}, \mathbf{Q}_1^{\mathrm{D}}, \ldots, \boldsymbol{s}_K^{\mathrm{U}}, \mathbf{Q}_K^{\mathrm{D}}) = \omega_n^{\mathrm{D}} \sum_k \times b\big(\gamma_{k,n}^{\mathrm{D}}(\boldsymbol{s}_k^{\mathrm{U}}, \mathbf{Q}_k^{\mathrm{D}}, \sigma_{k,n}^{\mathrm{D}})\big) \tag{39b}$$

The coordinate block associated with f_n^{U} is $[s_{1,n}^{\mathrm{U}}, \ldots, s_{K,n}^{\mathrm{U}}]^{\mathcal{T}}$ and the one associated with f_n^{D} is $[\mathbf{Q}_{1,n}^{\mathrm{D}\,\mathcal{T}}, \ldots, \mathbf{Q}_{K,n}^{\mathrm{D}\,\mathcal{T}}]^{\mathcal{T}}$. For ease of notation, define $f_{k,n}^{\mathrm{U}}(\boldsymbol{s}_k^{\mathrm{U}}) = \omega_n^{\mathrm{U}} b\big(\gamma_{k,n}^{\mathrm{U}}(\boldsymbol{s}_k^{\mathrm{U}}, \boldsymbol{\Sigma}_k^{\mathrm{U}})\big)$, $f_{k,n}^{\mathrm{D}}(\boldsymbol{s}_k^{\mathrm{U}}, \mathbf{Q}_k^{\mathrm{D}}) = \omega_n^{\mathrm{D}} b\big(\gamma_{k,n}^{\mathrm{D}}(\boldsymbol{s}_k^{\mathrm{U}}, \mathbf{Q}_k^{\mathrm{D}}, \sigma_{k,n}^{\mathrm{D}})\big)$, as well as the corresponding surrogate functions $\tilde{f}_{k,n}^{\mathrm{U}}(s_{k,n}^{\mathrm{U}}; \bar{\boldsymbol{s}}_{k,n}^{\mathrm{U}}, \bar{\mathbf{Q}}_k^{\mathrm{D}})$ and $\tilde{f}_{k,n}^{\mathrm{D}}(\mathbf{Q}_{k,n}^{\mathrm{D}}; \bar{\boldsymbol{s}}_{k,n}^{\mathrm{D}}, \bar{\mathbf{Q}}_k^{\mathrm{D}})$. From [42], it is known that functions f_n^{U} and f_n^{D} are concave in their associated coordinate blocks and convex in the other coordinate blocks. The gradients in the resulting surrogate functions are given by $\boldsymbol{a}_n^{\mathrm{U}\,\mathcal{T}} = [a_{1,n}^{\mathrm{U}}, \ldots, a_{K,n}^{\mathrm{U}}]$ and $\mathbf{A}_n^{\mathrm{D}\,\mathcal{T}} = [\mathbf{A}_{1,n}^{\mathrm{D}\,\mathcal{T}}, \ldots, \mathbf{A}_{K,n}^{\mathrm{D}\,\mathcal{T}}]$ with

$$a_{k,n}^{\mathrm{U}} = \sum_{m<n} \boldsymbol{h}_{k,n}^{\mathrm{U}\,\mathcal{H}} \boldsymbol{\Phi}_{k,m}^{\mathrm{U}} \boldsymbol{h}_{k,n}^{\mathrm{U}} + \sum_m \phi_{k,m}^{\mathrm{D}} h_{k,mn}^{\mathrm{XX}} \tag{40a}$$

$$\mathbf{A}_{k,n}^{\mathrm{D}} = \sum_{m>n} \phi_{k,m}^{\mathrm{D}} \boldsymbol{h}_{k,m}^{\mathrm{D}} \boldsymbol{h}_{k,m}^{\mathrm{D}\,\mathcal{H}}, \tag{40b}$$

where $\boldsymbol{\Phi}_{k,m}^{\mathrm{U}}$ is defined as in (35a) and where $\phi_{k,m}^{\mathrm{D}}$ expresses the sensitivity of the weighted bit loading for user m to changes in its interference-plus-noise power $\psi_{k,m}^{\mathrm{D}}$.

$$\phi_{k,n}^{\mathrm{D}} = \log(2)^{-1} \omega_n^{\mathrm{D}} \frac{\boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}} \bar{\mathbf{Q}}_{k,n}^{\mathrm{D}} \boldsymbol{h}_{k,n}^{\mathrm{D}}}{\bar{\psi}_{k,n}^{\mathrm{D}} (\Gamma \bar{\psi}_{k,n}^{\mathrm{D}} + \boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}} \bar{\mathbf{Q}}_{k,n}^{\mathrm{D}} \boldsymbol{h}_{k,n}^{\mathrm{D}})} \tag{41}$$

When a linear transceiver structure is considered, the summations in (40) should be over $\mathcal{N} \setminus \{n\}$, instead of $m < n$ or $m > n$.

When functions f_i are identified as in (39), the WSR maximization problem (13) does not fit the DSB problem structure from (30) due to the feasible space of the DS transmit covariance matrices being non-separable. As a result, the convergence results from [36]–[38] no longer apply. Nevertheless, convergence results for the non-separable case have recently become available [39], [43] and, more importantly, the DSB algorithm still leads to a simple solution strategy.

Despite the surrogate problem being non-separable across DS users, it is still separable across US users, i.e. the surrogate problem can be independently solved for each US user. The per user US surrogate problems can be further decoupled over tones by dual decomposition. The resulting US Lagrange dual problem

$$\underset{\lambda_n^{\mathrm{U}} \in \mathbb{R}_+}{\text{minimize}}\ q_n^{\mathrm{U}}(\lambda_n^{\mathrm{U}}) \tag{42}$$

is one-dimensional and convex, such that a simple bisection search algorithm can be used. As the US surrogate problem is convex and a primal feasible point exists, the duality gap between the surrogate problem and (42) is zero [30, Section 5.2.3], removing the necessity of invoking the time sharing property to have a zero duality gap. The Lagrange dual function is defined as

$$q_n^{\mathrm{U}}(\lambda_n^{\mathrm{U}}) = \sum_k \max_{s_{k,n}^{\mathrm{U}}} \left\{ \tilde{f}_{k,n}^{\mathrm{U}}(s_{k,n}^{\mathrm{U}}; \bar{\boldsymbol{s}}_k^{\mathrm{U}}, \bar{\mathbf{Q}}_k^{\mathrm{D}}) - \lambda_n^{\mathrm{U}} s_{k,n}^{\mathrm{U}} \right\} + \sum_n \lambda_n^{\mathrm{U}} P_{\mathrm{tot}}^{\mathrm{U}}, \tag{43}$$

The optimal $s_{k,n}^{\mathrm{U}}$ is given by (36a) with $a_{k,n}^{\mathrm{U}}$ calculated as in (40a) instead of (34a) [25]. As the duality gap is zero, the solution to the US surrogate problem is given by the unique maximizer in (43) for the optimal $\boldsymbol{\lambda}^{\mathrm{D}}$ from (42) [30, Section 5.5.5].

The surrogate problem is non-separable across DS users. However, the Lagrange dual decomposition decouples the problem not only across tones, but also across users. The DS Lagrange dual problem is given by

$$\underset{\boldsymbol{\lambda}^{\mathrm{D}} \in \mathbb{R}_+^N}{\text{minimize}}\ q^{\mathrm{D}}(\boldsymbol{\lambda}^{\mathrm{D}}), \tag{44}$$

where the Lagrange dual function $q^{\mathrm{D}}(\boldsymbol{\lambda}^{\mathrm{D}})$ is defined as

$$q^{\mathrm{D}}(\boldsymbol{\lambda}^{\mathrm{D}}) = \mathbf{1}_N^{\mathcal{T}} \boldsymbol{\lambda}^{\mathrm{D}} P_{\mathrm{tot}}^{\mathrm{D}} + \sum_k \sum_n \max_{\substack{\mathbf{Q}_{k,n}^{\mathrm{D}} \succeq 0 \\ \text{Hermitian}}} \left\{ \tilde{f}_{k,n}^{\mathrm{D}}(\mathbf{Q}_{k,n}^{\mathrm{D}}; \bar{\boldsymbol{s}}_k^{\mathrm{D}}, \bar{\mathbf{Q}}_k^{\mathrm{D}}) - \mathrm{tr}(\boldsymbol{\Lambda}^{\mathrm{D}} \mathbf{Q}_{k,n}^{\mathrm{D}}) \right\}. \tag{45}$$

As before, a closed form expression for the per tone per user maximizer in (45) is available [42], i.e.

$$\boldsymbol{t}_{k,n}^{\mathrm{D}} = (\mathbf{A}_{k,n}^{\mathrm{D}} + \boldsymbol{\Lambda}^{\mathrm{D}})^{-1}\boldsymbol{h}_{k,n}^{\mathrm{D}} \tag{46a}$$

$$s_{k,n}^{\mathrm{D}} = \left[\frac{\omega_n^{\mathrm{D}}\log(2)^{-1} - \Gamma\bar{\psi}_{k,n}^{\mathrm{D}}(\boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}}\boldsymbol{t}_{k,n}^{\mathrm{D}})^{-1}}{\boldsymbol{h}_{k,n}^{\mathrm{D}\,\mathcal{H}}\boldsymbol{t}_{k,n}^{\mathrm{D}}}\right]^{+} \tag{46b}$$

$$\mathbf{Q}_{k,n}^{\mathrm{D}} = s_{k,n}^{\mathrm{D}}\boldsymbol{t}_{k,n}^{\mathrm{D}}\boldsymbol{t}_{k,n}^{\mathrm{D}\,\mathcal{H}}. \tag{46c}$$

What remains is to solve the dual problem (44). As before, one possibility is to use a subgradient scheme. The dual function $q^{\mathrm{D}}(\boldsymbol{\lambda}^{\mathrm{D}})$ is however not only convex, but also continuously differentiable on its domain (see Appendix A, Proposition 5). More efficient gradient descent schemes can therefore be employed, constituting a vast improvement over the subgradient method.

The DS dual problem can be further simplified by removing the positivity constraint on $\boldsymbol{\lambda}^{\mathrm{D}}$, which is justified by observing that any solution of the DS surrogate problem must satisfy the total power constraints with equality. Indeed, for any DS resource allocation with $P_n^{\mathrm{D}} < P_{\mathrm{tot}}^{\mathrm{D}}$, the objective function value of the surrogate problem can be increased by adding $\xi\boldsymbol{e}_n\boldsymbol{e}_n^{\mathcal{H}}$ to $\mathbf{Q}_{k,N}^{\mathrm{D}}$, with $\xi > 0$.[5] The inequality constraints in the DS surrogate problem can thus be replaced with equality constraints, resulting in an unconstrained dual problem. As such, methods for unconstrained optimization can be used. We opt for a simple gradient descent algorithm with Armijo backtracking, which, assuming the starting point $\boldsymbol{\lambda}^{\mathrm{D}}$ is dual feasible and given the fact that $q(\boldsymbol{\lambda}^{\mathrm{D}})$ is continuously differentiable on its domain, is guaranteed to converge to the global optimum of (44) [44].

The resulting algorithm, referred to as full-duplex primal domain distributed spectrum balancing (FDX-PD-DSB), is summarized in Algorithm 3. The complexity of Algorithm 3 is determined by the calculation of $\mathbf{Q}_{k,n}^{\mathrm{D}}$ $\forall k, n$ in line 26, which involves determining a matrix inverse (46a) resulting in a complexity of $\mathcal{O}(N^3)$ for each $\mathbf{Q}_{k,n}^{\mathrm{D}}$. However, by using the Sherman-Morrison identity [41], $(\mathbf{A}_{k,n-1}^{\mathrm{D}} + \boldsymbol{\Lambda}^{\mathrm{D}})^{-1}$ can be obtained from $(\mathbf{A}_{k,n}^{\mathrm{D}} + \boldsymbol{\Lambda}^{\mathrm{D}})^{-1}$ with a complexity of only $\mathcal{O}(N^2)$. All other calculations involved in line 26 have a per user per tone complexity of at most $\mathcal{O}(N^2)$. Each iteration of line 26 therefore results in a complexity of $\mathcal{O}(KN^3)$. With I_{DSB} the number of iterations required by the DSB algorithm to converge and I_{GD} the number of iterations required by the gradient descent algorithm to converge, the resulting total complexity of FDX-PD-DSB is $\mathcal{O}(I_{\mathrm{DSB}}I_{\mathrm{GD}}KN^3)$.

In order to establish convergence guarantees for the algorithms of [39], [43], assumptions are made that are similar to **C.1**-**C.5**. The surrogates of FDX-PD-DSB do however not satisfy **C.5**, such that the results in [39], [43] do not warrant convergence of the primal variables $s_{k,n}^{\mathrm{U}}$ and $\mathbf{Q}_{k,n}^{\mathrm{D}}$.[6] All simulations have however shown monotonically increasing WSR values that converge after only a few iterations.

Algorithm 3 FDX-PD-DSB

1: Choose $\bar{\boldsymbol{s}}_k^{\mathrm{U}} \in \mathbb{R}_+^N$, $\bar{\boldsymbol{s}}_k^{\mathrm{D}} \in \mathbb{R}_+^N$, and set $\bar{\boldsymbol{T}}_k^{\mathrm{D}} \leftarrow \mathbf{H}_k^{\mathrm{D}\,-\mathcal{H}}$ $\forall k$
2: Set $\bar{\mathbf{Q}}_{k,n}^{\mathrm{D}} \leftarrow \bar{s}_{k,n}^{\mathrm{D}}\bar{\boldsymbol{t}}_{k,n}^{\mathrm{D}}\bar{\boldsymbol{t}}_{k,n}^{\mathrm{D}\,\mathcal{H}}$ $\forall k, n$
3: Choose $\epsilon_\tau \in (0, 1)$, and set $\tau^+ \leftarrow 1$, $\tau^- \leftarrow 1 - \tau^+$
4: **while** $\sum_n \omega_n^{\mathrm{U}} R_n^{\mathrm{U}} + \sum_n \omega_n^{\mathrm{D}} R_n^{\mathrm{D}}$ not converged **do**
5: Calculate $a_{k,n}^{\mathrm{U}}$ and $\mathbf{A}_{k,n}^{\mathrm{D}}$ $\forall k, n$ using (40)
6: $\boldsymbol{s}_n^{\mathrm{U}} \leftarrow \mathrm{mUS}(\boldsymbol{a}_n^{\mathrm{U}}, \bar{\boldsymbol{\Psi}}_n^{\mathrm{U}})$ $\forall n$
7: $\{\mathbf{Q}_1^{\mathrm{D}}, \ldots, \mathbf{Q}_N^{\mathrm{D}}\} \leftarrow \mathrm{mDS}(\mathbf{A}_1^{\mathrm{D}}, \ldots, \mathbf{A}_N^{\mathrm{D}}, \bar{\boldsymbol{\psi}}_1^{\mathrm{D}}, \ldots, \bar{\boldsymbol{\psi}}_N^{\mathrm{D}})$
8: Decompose $\mathbf{Q}_{k,n}^{\mathrm{D}}$ into $s_{k,n}^{\mathrm{D}}$ and $\boldsymbol{t}_{k,n}^{\mathrm{D}}$ $\forall k, n$
9: Set $\bar{\boldsymbol{s}}_n^{\mathrm{U}} \leftarrow \tau^-\bar{\boldsymbol{s}}_n^{\mathrm{U}} + \tau^+\boldsymbol{s}_n^{\mathrm{U}}$, $\bar{\boldsymbol{s}}_n^{\mathrm{D}} \leftarrow \tau^-\bar{\boldsymbol{s}}_n^{\mathrm{D}} + \tau^+\boldsymbol{s}_n^{\mathrm{D}}$ $\forall n$
10: Set $\bar{\boldsymbol{t}}_{k,n}^{\mathrm{D}} \leftarrow \tau^-\bar{\boldsymbol{t}}_{k,n}^{\mathrm{D}} + \tau^+\boldsymbol{t}_{k,n}^{\mathrm{D}}$ $\forall k, n$
11: Set $\tau^+ \leftarrow \tau^+(1 - \tau^+\epsilon_\tau)$, $\tau^- \leftarrow 1 - \tau^+$
12: **function** $\mathrm{mUS}(\boldsymbol{a}_n^{\mathrm{U}}, \bar{\boldsymbol{\Psi}}_n^{\mathrm{U}})$
13: Choose search interval $[\lambda_{\min}^{\mathrm{U}}, \lambda_{\max}^{\mathrm{U}}] \subset \mathbb{R}_+$
14: **while** $|\sum_k s_{k,n}^{\mathrm{U}} - P_{\mathrm{tot}}^{\mathrm{U}}| > \delta_{\mathrm{P}}$ & $\lambda_n^{\mathrm{U}} > \delta_\lambda$ **do**
15: Set $\lambda^{\mathrm{U}} \leftarrow (\lambda_{\min}^{\mathrm{U}} + \lambda_{\max}^{\mathrm{U}})/2$
16: Calculate $s_{k,n}^{\mathrm{U}}$ $\forall k$ for λ^{U} using (36a)
17: **if** $P_n^{\mathrm{U}} > P_{\mathrm{tot}}^{\mathrm{U}}$ **then** $\lambda_{\min}^{\mathrm{U}} \leftarrow \lambda^{\mathrm{U}}$ **else** $\lambda_{\max}^{\mathrm{U}} \leftarrow \lambda^{\mathrm{U}}$
18: **function** $\mathrm{mDS}(\mathbf{A}_1^{\mathrm{D}}, \ldots, \mathbf{A}_N^{\mathrm{D}}, \bar{\boldsymbol{\psi}}_1^{\mathrm{D}}, \ldots, \bar{\boldsymbol{\psi}}_N^{\mathrm{D}})$
19: Choose $[\boldsymbol{g}_{\mathrm{prev}}]_n > \delta_{\mathrm{P}}$ $\forall n$, $\boldsymbol{\lambda}^{\mathrm{D}} \in \mathbb{R}_+^N$, $\mu_{\max} \gg 0$
20: Set $q_{\mathrm{prev}} \leftarrow q^{\mathrm{D}}(\boldsymbol{\lambda}^{\mathrm{D}})$ using (45), $\mathbf{Q}_{k,n}^{\mathrm{D}}$ $\forall k, n$ from (46)
21: **while** $|[\boldsymbol{g}_{\mathrm{prev}}]_n| > \delta_{\mathrm{P}}$ for any n **do**
22: Set $\boldsymbol{g}_{\mathrm{prev}} \leftarrow P_{\mathrm{tot}}^{\mathrm{D}}\mathbf{1}_N - \sum_k \sum_n \mathrm{diag}\{\mathbf{Q}_{k,n}^{\mathrm{D}}\}$
23: Set $\mu^{\mathrm{D}} \leftarrow \mu_{\max}$, $q_{\mathrm{new}} \leftarrow q_{\mathrm{prev}}$
24: **while** $q_{\mathrm{new}} > q_{\mathrm{prev}} - \mu_{\mathrm{Armijo}}\mu^{\mathrm{D}}\|\boldsymbol{g}_{\mathrm{prev}}\|_2^2$ **do**
25: Set $\mu^{\mathrm{D}} \leftarrow \mu^{\mathrm{D}}/2$ and $\Delta\boldsymbol{\lambda}^{\mathrm{D}} \leftarrow -\mu^{\mathrm{D}}\boldsymbol{g}_{\mathrm{prev}}$
26: Set $q_{\mathrm{new}} \leftarrow q^{\mathrm{D}}(\boldsymbol{\lambda}^{\mathrm{D}} + \Delta\boldsymbol{\lambda}^{\mathrm{D}})$ using (45), $\mathbf{Q}_{k,n}^{\mathrm{D}}$ $\forall k, n$ from (46)
27: Set $q_{\mathrm{prev}} \leftarrow q_{\mathrm{new}}$, $\boldsymbol{\lambda}^{\mathrm{D}} \leftarrow \boldsymbol{\lambda}^{\mathrm{D}} + \Delta\boldsymbol{\lambda}^{\mathrm{D}}$

In the derivation of FDX-PD-DSB, it has already been indicated how to adapt Algorithm 3 when a linear transceiver structure is considered. Other versions of FDX-PD-DSB can be obtained by including bit cap or spectral mask constraints. Bit cap constraints are readily enforced by using the augmented bit loading function of Section V-B. Spectral mask constraints can be enforced as for FDX-DD-DSB, i.e. by replacing $[\cdot]^+$ in (36a) with $[\cdot]_0^{s_k^{\mathrm{mask}}}$ in US and by dualizing the spectral mask constraints in DS. It is noted that, due to the convexity of the surrogate problems, the duality gap for DS will still be zero, such that dualizing the DS spectral mask constraints will not lead to problematic convergence behavior. Finally, problem (13) can be altered by relaxing the assumption of perfect DPU-side NEXT cancellation. One can readily adapt FDX-PD-DSB to networks with imperfect DPU-side NEXT cancellation by adding a term to (40b) modeling the DPU-side NEXT influence. As such,

[5] This argument assumes $h_{k,nN}^{\mathrm{D}} \neq 0$. If there does not exist any tone k for which this is true, the same argument can be repeated for user $N-1$, as $h_{k,nN}^{\mathrm{D}} = 0$ and (40b) imply $\boldsymbol{e}_n \in \mathrm{null}(\mathbf{A}_{k,N-1}^{\mathrm{D}})$.

[6] Even though strong convexity cannot be established, the results in Appendix A do imply that the surrogate problems have unique solution.

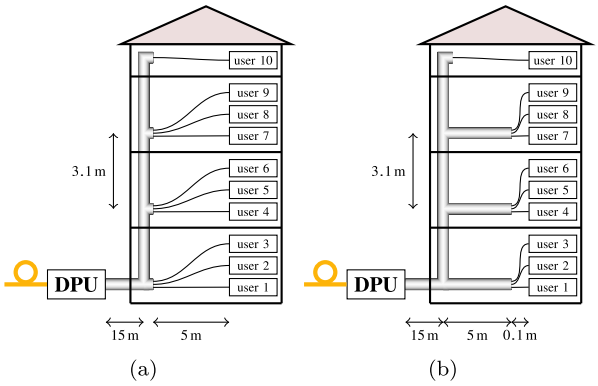

FIGURE 2. A DSL network in a 10-user fiber to the building scenario, which is considered to be the main use case for G.mgfast.

FDX-PD-DSB can be applied to the resource allocation problems of wireless FDX systems [13]–[17] as well. It is noted that this is in contrast with FDX-OSB and FDX-DD-DSB, which were both founded on the MAC-BC duality theory from Section IV that required perfect DPU-side NEXT cancellation.

Another important property of FDX-PD-DSB is that after each outer iteration, a new feasible resource allocation is provided. Therefore, FDX-PD-DSB is a so-called real-time DSM algorithm, as defined in [45]. From FDX-PD-DSB, other low complexity real-time DSM algorithms for FDX DSL networks can be derived. For example, an iterative waterfilling (IWF) type algorithm [46] can be obtained by setting $a_{k,n}^{\mathrm{U}}$ and $\mathbf{A}_{k,n}^{\mathrm{D}}$ to 0. In addition, constant offset autonomous spectrum balancing (CO-ASB) type algorithms [47], [48] can be obtained by setting $a_{k,n}^{\mathrm{U}}$ and $\mathbf{A}_{k,n}^{\mathrm{D}}$ to a predetermined constant value.

An important difference between the proposed SCA method, and the state-of-the-art (SoA) SCA methods in [15]–[17] is highlighted. Assuming the multi-carrier extensions of the SCA algorithms in [15]–[17] also entail solving the Lagrange dual of their respective surrogate problems, the SoA SCA methods rely on iterative solvers to deal with the evaluation of their respective Lagrange dual functions. This is in stark contrast with the evaluation the Lagrange dual function in FDX-PD-DSB, which can be done analytically and has a computational complexity that is mainly determined by a single matrix inversion in (46a).

VI. SIMULATION RESULTS

In this section the performance of the proposed DSM algorithms, as well as that of FDX transmission itself, is analyzed for the two DSL networks illustrated in Fig. 2. Each network consists of a fiber to the building (FTTB) scenario with $N = 10$. Such FTTB scenarios are considered to be the main use case for G.mgfast. The channel matrices for these scenarios are modeled based on lab measurements. The main difference between the considered scenarios is the separation length, i.e. the length of the copper line separating two NTs (e.g. the separation length of user 1 and user 2 in Fig. 2a is 10 m). Due to the larger separation length in scenario (a), the NEXT signals travel a longer distance before the copper lines meet, such that more attenuation has taken place when interference sets in.

As ITU-T recommendations for G.mgfast are not yet publicly available, the employed system model parameters are chosen conform G.fast recommendations [8], [49]. The maximum total per line power is $P_{\mathrm{tot}}^{\mathrm{U}} = P_{\mathrm{tot}}^{\mathrm{D}} = 4$ dBm, and spectral mask constraints range from -65 dBm/Hz to -79 dBm/Hz. The tone spacing, symbol rate, and SNR-gap are $\Delta_{\mathrm{f}} = 51.75$ kHz, $f_{\mathrm{s}} = 48$ kHz, and $\Gamma = 10$ dB. A bit cap of $b_{\max} = 15$ is assumed. The number of tones is $K = 4096$, corresponding to a bandwidth of up to 212 MHz. In addition to perfect DPU-side echo cancellation, perfect per user NT-side echo cancellation is assumed.

A. DSB OPTIMALITY GAP

The performance differences between the OSB and DSB algorithms are analyzed for a network consisting of the first three users of Fig. 2b. All weights are chosen as $\omega_n^{\mathrm{U}} = \omega_n^{\mathrm{D}} = 1$, and no bit cap or spectral mask constraints are taken into account. A linear transceiver structure is considered, which in simulations is often seen to result in the existence of multiple stationary points in the objective function. A system with a linear transceiver structure is therefore more adversarial for the DSB algorithms than one using a non-linear transceiver structure. Hence, it will highlight the performance difference between the OSB and DSB algorithms.

Executing FDX-OSB for a 3-user FDX DSL network is no easy feat. To achieve convergence in an acceptable amount of time, the dual function (15) is evaluated using 7 parallel computational threads, each employing GPU acceleration to evaluate line 14 of Algorithm 1. Using this setup, a single evaluation of the dual function (15) takes about 3 min.

The bit loading resulting from FDX-OSB, FDX-DD-DSB, and FDX-PD-DSB are shown in Fig. 3, and the achieved per user data rates are given in Table 1. Both Fig. 3 and Table 1 demonstrate that the solutions found by the DSB algorithms are very similar to the solution found by the OSB algorithm. The WSR obtained from the DSB algorithms is even slightly higher than the WSR obtained from the OSB algorithm. This is explained by the discrete grid search of FDX-OSB, which, due to its finite granularity, can result in a slightly suboptimal solution when compared to the DSB algorithms.

B. CONVERGENCE OF THE DSB ALGORITHMS

The convergence of the proposed DSB algorithms is analyzed for the 10-user FTTB network of Fig. 2a. Spectral mask and bit cap constraints are considered to be inactive. For various DS weights (with $\omega_n^{\mathrm{U}} = 1 - \omega_n^{\mathrm{D}}$), the total number of iterations I and the final WSR value are shown in Fig. 4. For FDX-PD-DSB, I equals the number of outer DSB updates times the number of gradient updates, i.e.

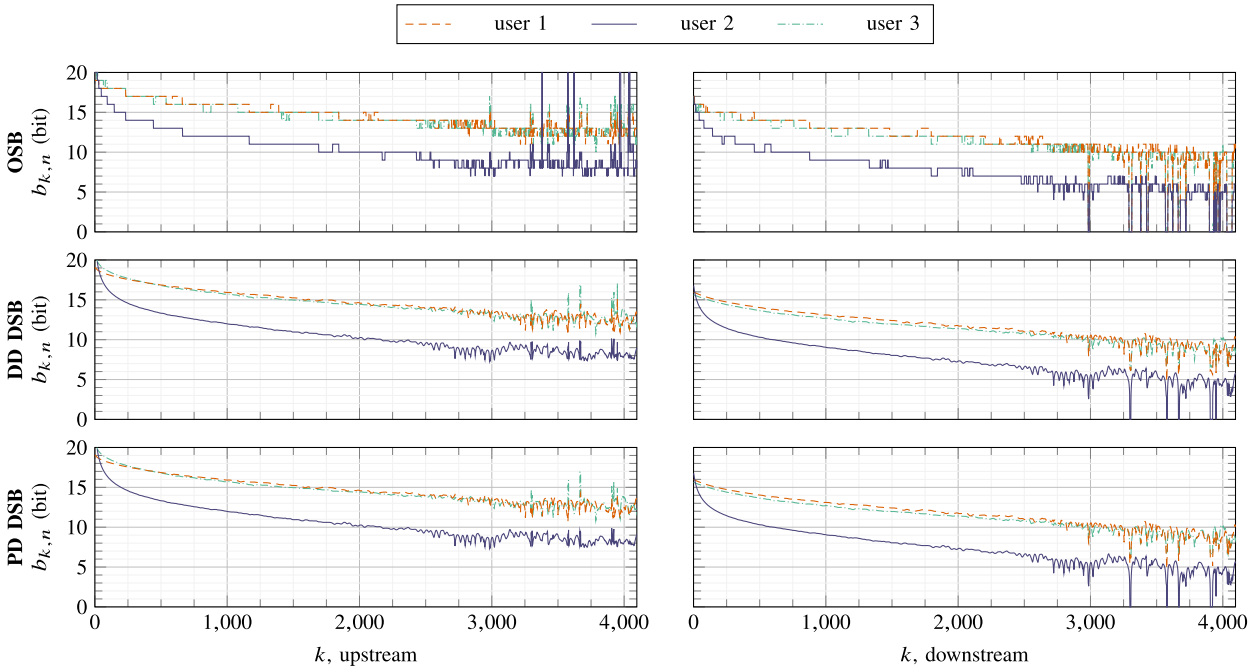

FIGURE 3. Illustration of the optimality of DSB algorithms. It is seen that the solution obtained by DSB algorithms closely resembles the one obtained by the OSB algorithm.

TABLE 1. Data rate and WSR for a three user FDX DSL network (Gbit/s).

	FDX-OSB			FDX-DD-DSB			FDX-PD-DSB		
	user 1	user 2	user 3	user 1	user 2	user 3	user 1	user 2	user 3
US	2.87	2.08	2.85	2.89	2.09	2.87	2.89	2.09	2.88
DS	2.27	1.49	2.21	2.30	1.47	2.24	2.29	1.47	2.23
WSR	13.78			13.86			13.86		

$I = I_{\text{DSB}} I_{\text{GD}}$, whereas for FDX-DD-DSB, it equals the number of subgradient updates times the average number of inner DSB iterations i.e. $I = I_{\text{SG}} I_{\text{DSB}}$. In the stop criteria of the Lagrange multiplier searches, the power constraints should be satisfied with an accuracy of $\delta_{\text{P}} = 10^{-3} P^{\text{U}}_{\text{tot}} = 10^{-3} P^{\text{D}}_{\text{tot}}$. The DSB algorithms are stopped when the relative objective function increase between two consecutive iterations drops below 10^{-3}. FDX-PD-DSB is initialized with non-linear QRD-based zero-forcing precoder matrices and a flat spectrum satisfying all power constraints, and FDX-DD-DSB is initialized by setting all symbol powers equal to zero.

The total number of iterations and final WSR values are given in Fig. 4. For each DSB algorithm, three modes of operation are considered, which differ in the way the DSB algorithm updates its coordinate blocks in each iteration. The following update rules are considered.

- Data points labeled $\text{DD}_{\text{Jacobi}}$ and $\text{PD}_{\text{Jacobi}}$ are obtained by executing FDX-DD-DSB and FDX-PD-DSB with full Jacobi-style updates, as described in Algorithm 2 and Algorithm 3, with $\epsilon_\tau = 0$.
- Data points labeled $\text{DD}_{\text{part. Jacobi}}$ and $\text{PD}_{\text{part. Jacobi}}$ are obtained by letting FDX-DD-DSB and FDX-PD-DSB alternate between updating all US and all DS variables, with $\epsilon_\tau = 0$.

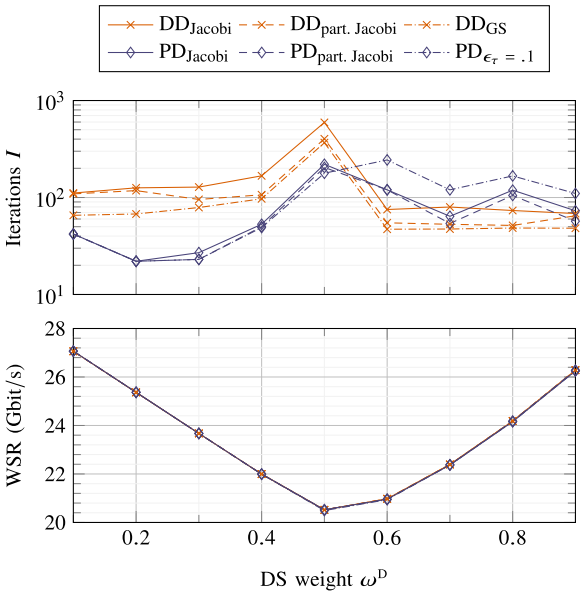

FIGURE 4. Number of iterations required for convergence and final weighted sum rate value, obtained by executing FDX-DD-DSB and FDX-PD-DSB for the network of Fig. 2a.

- Data points labeled DD_{GS} are obtained by updating the power variable of only a single user after each iteration of the DSB algorithm, with $\epsilon_\tau = 0$.

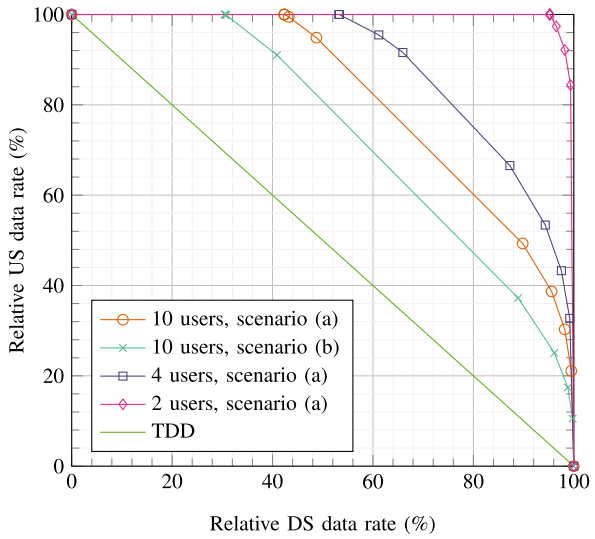

FIGURE 5. Average achievable US and DS rates of FDX transmission in the DSL networks of Fig. 2, relative to maximum achievable TDD rates.

- Data points labeled $\text{PD}_{\epsilon_\tau = .1}$ are obtained by executing FDX-PD-DSB with full Jacobi-style updates, as described in Algorithm 3 with $\epsilon_\tau = 0.1$.

In Fig. 4, a single iteration of the DSB algorithm, as counted in I, consists of updating the resource allocation variables of all users exactly once.

Fig. 4 shows that the final WSR obtained from FDX-DD-DSB and FDX-PD-DSB are very similar. The number of iterations required for convergence can however differ quite significantly. It should however be noted that, even though the per iteration complexity of the different algorithms is the same,[7] I is not a perfect indicator for the actual number of additions and multiplications that are executed by each algorithm. Indeed, in practical implementations, the execution time of FDX-PD-DSB is often seen to be in the order of tens of seconds, whereas executing FDX-DD-DSB usually requires several minutes.

It is seen that for $\omega^{\text{D}} > 0.5$, FDX-PD-DSB requires more iterations to converge than for $\omega^{\text{D}} < 0.5$. In these cases, it takes more iterations of the DSB algorithm to make the resource allocation converge on some of the tones. However, as each DSB iteration in FDX-PD-DSB updates all resource allocation variables, this increased number of DSB iterations has a significant impact on the execution time. By contrast, each DSB iteration in FDX-DD-DSB updates only the resource allocation of a single tone, such that slower convergence on some of the tones does not significantly affect the execution time of the entire algorithm.

Additional simulations have shown that FDX-PD-DSB on average requires far fewer iterations to solve a single dual problem than FDX-DD-DSB due to the employed gradient descent algorithm. An advantage of the gradient descent algorithm with backtracking is that it requires no parameter tuning. Indeed, convergence of the subgradient method is quite sensitive to its parameters, especially to the choice of the step size μ, and finding good values for them requires some experience.

The convergence of FDX-PD-DSB is also compared to that of the WMMSE algorithm. For this purpose, the WMMSE algorithm from [13] is adapted such that it can be applied to multi-carrier systems by using techniques from [12]. As the system model in [13] includes neither an SNR-gap nor non-linear transceivers, simulations will be executed for $\Gamma = 0$ dB and linear transceiver structures. The complexity of WMMSE is dominated by solving the Lagrange dual problem of the downstream precoding matrix optimization, where the most expensive operation in each iteration is a matrix inversion.

[7] This is not true for FDX-DD-DSB with Gauss Seidel updating, which has a per iteration complexity of $\mathcal{O}(N^4)$.

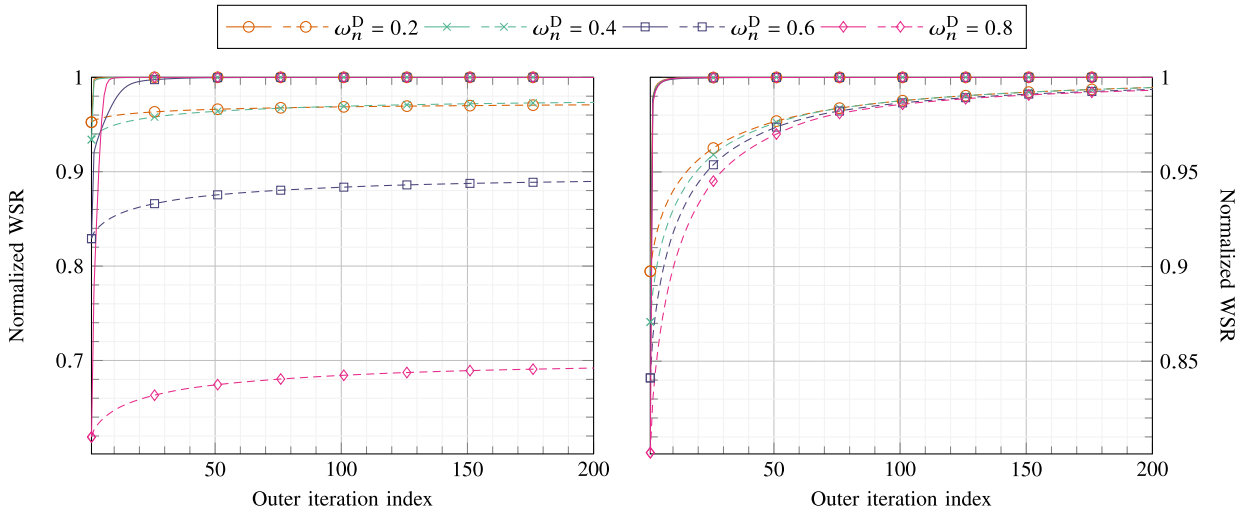

FIGURE 6. Convergence of the WMMSE and FDX-PD-DSB algorithms. Dashed curves correspond to WMMSE, solid curves to FDX-PD-DSB. Both algorithms are executed for the 10-user FTTB network of Fig. 2a. Both plots consider the same G.fast network, but are obtained using different values for the noise floor: $\sigma_{k,n} = -140$ dBm/Hz for the left plot, and $\sigma_{k,n} = -110$ dBm/Hz for the right plot.

As such, an outer iteration of WMMSE has the same asymptotic complexity as an outer iteration of FDX-PD-DSB.

The convergence experiments are performed for the 10-user FTTB network of Fig. 2a. FDX-PD-DSB is executed with $\epsilon_\tau = 0.01$. Both FDX-PD-DSB and WMMSE are initialized with linear ZF precoders and with the same randomly generated transmit powers. For various DS weights (with $\omega_n^{\mathrm{U}} = 1 - \omega_n^{\mathrm{D}}$), the obtained normalized[8] WSR values are given in the left plot of Fig. 6. It is observed that FDX-PD-DSB has converged after 50 iterations, whereas WMMSE has not yet converged after 200 iterations. In addition, a significant performance gap between the two algorithms is observed, especially for larger values of ω_n^{D}: after 200 iterations, FDX-PD-DSB outperforms WMMSE by 2% to 30%. These observations can be explained by the high SINR values at the optimum, a condition under which WMMSE is seen to suffer from a very slow convergence. When the SINR values are artificially decreased by raising the noise floor from −140 dBm/Hz to −110 dBm/Hz, as has been done to obtain convergence results in the right plot of Fig. 6, WMMSE is observed to converge faster. However, even in a setting more favorable to WMMSE, FDX-PD-DSB still requires far fewer iterations to obtain a high WSR.

C. PERFORMANCE OF MU FDX DSL NETWORKS

This section evaluates the FDX performance of DSL networks in the scenarios of Fig. 2. Both spectral mask and bit cap constraints are considered to be active. Fig. 5 shows the obtained results, consisting of the average achievable US and DS rates of FDX transmission, relative to the maximum achievable TDD rates.

The simulation reveals significant performance gains of FDX transmission over standard TDD transmission. In the 10-user networks of Fig. 2, for instance, at an average US rate of 50%, FDX transmission achieves a DS rate of 77.5% and 89% in scenarios (a) and (b) respectively, instead of the 50% rate achieved by TDD. The slightly higher gain in scenario (a) is due to the larger separation length between the NTs on each floor, resulting in more attenuated NT-side NEXT signals. Moreover, scenario (a) with only one active user per floor, i.e. $N = 4$, results in weaker overall NEXT, yielding even larger FDX gains (see Fig. 5). Even the extreme case with only 2 active users almost results in a doubled performance, close to the theoretical optimum. These results demonstrate that the level of NEXT interference at the NTs is *the* performance determining factor for MU FDX transmission in G.mgfast networks.

VII. CONCLUSION

Three novel DSM algorithms have been presented for the MU FDX DSL network. First, an OSB algorithm has been derived which calculates the globally optimal resource allocation, but does so at an exceedingly high computational cost. Two low-complexity DSB algorithms have then been proposed, for which simulations have shown that their performance is very close to what is achieved by the OSB algorithm. These DSB algorithms have then been used in performance simulations, which show that FDX transmission can indeed lead to significant performance gains in MU DSL networks.

[8]The normalization is such the best normalized WSR value found by either WMMSE or FDX-PD-DSB is equal to 1.

APPENDIX A
DIFFERENTIABILITY OF THE DUAL FUNCTION IN (45)

Differentiability of the dual function is strongly related to uniqueness of the maximizers in (45). It is therefore first proven that when $\mathbf{\Lambda}^{\mathrm{D}} \succ 0$, all maximization problems in (45) have a unique solution. Differentiability of $q(\boldsymbol{\lambda}^{\mathrm{D}})$ then follows from standard results in convex analysis.

Lemma 4 (Maximizer uniqueness): Optimization problem

$$\underset{\substack{\mathbf{Q}\succeq 0\\ \text{Hermitian}}}{\text{maximize}} \ \log\left(1 + \boldsymbol{h}^{\mathcal{H}}\mathbf{Q}\boldsymbol{h}\right) - \mathrm{tr}(\mathbf{Y}^{\mathcal{H}}\mathbf{Q}), \tag{47}$$

with $\boldsymbol{h} \neq \mathbf{0}_N$ and $\mathbf{Y}^{\mathcal{H}}$ positive definite Hermitian, has a unique solution, which furthermore is a rank-1 matrix.

Proof: The proof is based on the proof of [42, Proposition 1]. Let $\mathbf{L}\mathbf{L}^{\mathcal{H}}$ denote the Cholesky factorization of $\mathbf{Y}$. Furthermore, factorize $\mathbf{Q} = \mathbf{T}\mathbf{T}^{\mathcal{H}}$ where $\mathbf{T}$ is an $N \times N$ matrix. Defining $\tilde{\mathbf{T}} \triangleq \mathbf{L}^{\mathcal{H}}\mathbf{T}$, the following problem is equivalent to (47).

$$\underset{\tilde{\mathbf{T}}}{\text{maximize}} \ \log\left|\mathbf{I} + \tilde{\mathbf{T}}^{\mathcal{H}}\mathbf{L}^{-1}\boldsymbol{h}\boldsymbol{h}^{\mathcal{H}}\mathbf{L}^{-\mathcal{H}}\tilde{\mathbf{T}}\right| - \mathrm{tr}(\tilde{\mathbf{T}}\tilde{\mathbf{T}}^{\mathcal{H}})$$

Let $\mathbf{U}\mathbf{\Sigma}\mathbf{U}^{\mathcal{H}}$ denote the eigendecomposition of the Hermitian matrix $\mathbf{L}^{-1}\boldsymbol{h}\boldsymbol{h}^{\mathcal{H}}\mathbf{L}^{-\mathcal{H}}$. After defining $\hat{\mathbf{T}} \triangleq \mathbf{U}^{\mathcal{H}}\tilde{\mathbf{T}}$, the following problem is obtained, which is still equivalent to (47).

$$\underset{\hat{\mathbf{T}}}{\text{maximize}} \ \log\left|\mathbf{I} + \hat{\mathbf{T}}^{\mathcal{H}}\mathbf{\Sigma}\hat{\mathbf{T}}\right| - \mathrm{tr}(\hat{\mathbf{T}}\hat{\mathbf{T}}^{\mathcal{H}})$$

As $\mathbf{\Sigma}$ contains only a single non-zero element $[\mathbf{\Sigma}]_{ii} = \sigma$, the problem (47) simplifies to

$$\underset{\hat{\mathbf{T}}}{\text{maximize}} \ \log\left(1 + \sigma\|[\hat{\mathbf{T}}]_{\mathrm{row}i}\|_2^2\right) - \|\hat{\mathbf{T}}\|_F^2. \tag{48}$$

Equation (48) implies that any optimal $\hat{\mathbf{T}}$ has non-zero elements only in its ith row, as any non-zero element in $\hat{\mathbf{T}}$ that is not on the i-th row decreases the objective function value of (48). Therefore, problem (48) simplifies to

$$\underset{[\hat{\mathbf{T}}]_{\mathrm{row}i}}{\text{maximize}} \ \log\left(1 + \sigma\|[\hat{\mathbf{T}}]_{\mathrm{row}i}\|_2^2\right) - \|[\hat{\mathbf{T}}]_{\mathrm{row}i}\|_2^2,$$

Observe that any optimal $[\hat{\mathbf{T}}]_{\mathrm{row}i}$ must satisfy $\|[\hat{\mathbf{T}}]_{\mathrm{row}i}\|_2^2 = [1 - \sigma^{-1}]^+$. Therefore, any optimal $\hat{\mathbf{T}}$ yields

$$\mathbf{Q} = \mathbf{L}^{-\mathcal{H}}\mathbf{U}\mathbf{D}\mathbf{U}^{\mathcal{H}}\mathbf{L}^{-1}$$

with $\mathbf{D}$ containing only a single non-zero element $[\mathbf{D}]_{ii} = [1 - \sigma^{-1}]^+$. The solution to (47) is thus unique and rank-1. □

Proposition 5 (Continuous differentiability q^D): If $\boldsymbol{h}_{k,n}^D \neq \mathbf{0}_N \ \forall k, n$, then the dual function $q^D(\boldsymbol{\lambda}^D)$ defined by (45) is continuously differentiable on its domain.

Proof: First, it is argued that the domain of q^{D} is given by $\{\boldsymbol{\lambda}^{\mathrm{D}} \in \mathbb{R}_+^N \mid \lambda_n^{\mathrm{D}} > 0\,\forall n\}$. Then it is shown that on its domain, q^{D} is continuously differentiable.

If any $\lambda_n^{\mathrm{D}} \leq 0$, then the objective of the maximization in (45) can be made arbitrarily large by choosing $\mathbf{Q}_{k,N}^{\mathrm{D}} = \Omega \boldsymbol{e}_n \boldsymbol{e}_n{}^{\mathcal{H}}$, with $\Omega \to +\infty$.[9] The objective of the maximization in (45) is therefore unbounded such that $\boldsymbol{\lambda}^{\mathrm{D}}$ is outside domain of q^{D}. For any $\boldsymbol{\lambda}^{\mathrm{D}} \in \{\boldsymbol{\lambda}^{\mathrm{D}} \in \mathbb{R}_+^N \mid \lambda_n^{\mathrm{D}} > 0\,\forall n\}$ however, Lemma 4 implies that a finite maximizer exists, such that $q^{\mathrm{D}}(\boldsymbol{\lambda}^{\mathrm{D}})$ is finite also.

To prove that q^{D} is continuously differentiable on its domain, introduce

$$F_{k,n}(\mathbf{Q}_{k,n}^{\mathrm{D}}) = \begin{cases} -\tilde{f}_{k,n}^{\mathrm{D}}(\mathbf{Q}_{k,n}^{\mathrm{D}}; \bar{\boldsymbol{s}}_k^{\mathrm{D}}, \bar{\mathbf{Q}}_k^{\mathrm{D}}) & \text{if } \mathbf{Q}_{k,n}^{\mathrm{D}} \succeq 0, \\ & \text{and } \mathbf{Q}_{k,n}^{\mathrm{D}} = (\mathbf{Q}_{k,n}^{\mathrm{D}})^{\mathcal{H}} \\ +\infty & \text{otherwise} \end{cases}$$

and its conjugate function [50, Chapter 11]

$$F_{k,n}^*(\mathbf{Y}) \triangleq \sup_{\mathbf{Q}_{k,n}^{\mathrm{D}}} \left\{ \mathrm{tr}(\mathbf{Y}^{\mathcal{H}} \mathbf{Q}_{k,n}^{\mathrm{D}}) - F_{k,n}(\mathbf{Q}_{k,n}^{\mathrm{D}}) \right\}.$$

Using this definition of $F_{k,n}^*(\mathbf{Y})$, (45) can be rewritten as

$$q^{\mathrm{D}}(\boldsymbol{\lambda}^{\mathrm{D}}) = \sum_k \sum_n F_{k,n}^*\big(-\mathrm{diag}(\boldsymbol{\lambda}^{\mathrm{D}})\big) + \mathbf{1}_N{}^{\mathcal{T}} \boldsymbol{\lambda}^{\mathrm{D}} P_{\mathrm{tot}}^{\mathrm{D}}.$$

From [50, Theorem 11.8] it is known that the subdifferential of $F_{k,n}^*(\mathbf{Y})$ is given by the solution set of the optimization problem defining the conjugate function, i.e.

$$\partial F_{k,n}^*(\mathbf{Y}) = \arg\max_{\mathbf{Q}_{k,n}^{\mathrm{D}}} \left\{ \mathrm{tr}(\mathbf{Y}^{\mathcal{H}} \mathbf{Q}_{k,n}^{\mathrm{D}}) - F_{k,n}(\mathbf{Q}_{k,n}^{\mathrm{D}}) \right\}.$$

Lemma 4 implies that $\partial F_{k,n}^*(\mathbf{Y})$ is a singleton whenever $\mathbf{Y}$ is negative definite and Hermitian, i.e. $F_{k,n}^*(\mathbf{Y})$ is differentiable at any negative definite Hermitian $\mathbf{Y}$. Moreover, the proof of [50, Theorem 11.8] mentions that "$\partial F_{k,n}^*(\bar{\mathbf{Y}})$ is singleton if and only if $F_{k,n}^*(\mathbf{Y})$ is strictly differentiable at $\bar{\mathbf{Y}}$". $F_{k,n}^*(\mathbf{Y})$ is thus strictly differentiable on the open set of negative definite Hermitian matrices, implying that $F_{k,n}^*(\mathbf{Y})$ is continuously differentiable on the same set [50, Corollary 9.19]. It immediately follows that $q^{\mathrm{D}}(\boldsymbol{\lambda}^{\mathrm{D}})$ is continuously differentiable on its domain. □

ACKNOWLEDGMENT

The scientific responsibility is assumed by its authors. This paper was presented in part at the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, Montreal, Canada, October 2017 [1].

REFERENCES

[1] 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. IEEE 28th Annu. Int. Symp. Pers. Indoor Mobile Radio Commun.*, Montreal, QC, Canada, Oct. 2017, pp. 1–5.

[2] W. Coomans, R. B. Moraes, K. Hooghe, A. Duque, J. Galaro, M. Timmers, A. J. van Wijngaarden, M. Guenach, and J. Maes, "XG-fast: The 5th generation broadband," *IEEE Commun. Mag.*, vol. 53, no. 12, pp. 83–88, Dec. 2015.

[9] See footnote 5.

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

[4] *Asymmetric Digital Subscriber Line (ADSL) Transceivers*, document ITU-T G.992.1, International Telecommunication Union, ITU-T Study Group 15, 1999.

[5] *Asymmetric Digital Subscriber Line Transceivers 2 (ADSL2)*, document ITU-T G.992.3, International Telecommunication Union, ITU-T Study Group 15, 2002.

[6] *Very High Speed Digital Subscriber Line Transceivers*, document ITU-T G.993.1, International Telecommunication Union, ITU-T Study Group 15, 2004.

[7] *Very High Speed Digital Subscriber Line Transceivers 2 (VDSL2)*, document ITU-T G.993.2, International Telecommunication Union, ITU-T Study Group 15, 2007.

[8] *Fast Access to Subscriber Terminals (G.Fast)—Physical Layer Specification*, document ITU-T G.9701, International Telecommunication Union, ITU-T Study Group 15, 2015.

[9] A. R. Forouzan, M. Moonen, J. Maes, and M. Guenach, "Dynamic bandplanning for vectored DSL," *IEEE Trans. Commun.*, vol. 62, no. 1, pp. 302–315, Jan. 2014.

[10] R. Strobel, "Copper transmission for multi-gigabit hybrid copper-fiber access networks," in *Proc. Int. Workshop Fiber Opt. Access Netw.*, Munich, Germany, Nov. 2017, pp. 1–6.

[11] P. Tsiaflakis, Y. Lefevre, W. Coomans, and J. Maes, "Friendly full duplex: A multi-user full duplex method for MGfast in coexistence with G.fast," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Abu Dhabi, United Arab Emirates, Dec. 2018, pp. 1–6.

[12] R. B. Moraes, P. Tsiaflakis, J. Maes, and M. Moonen, "DMT MIMO IC rate maximization in DSL with per-transceiver power constraints," *Signal Process.*, vol. 101, pp. 87–98, Aug. 2014.

[13] S. Li, R. D. Murch, and V. K. N. Lau, "Linear transceiver design for full-duplex multi-user MIMO system," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Sydney, NSW, Australia, Jun. 2014, pp. 4921–4926.

[14] A. C. Cirik, R. Wang, Y. Hua, and M. Latva-Aho, "Weighted sum-rate maximization for full-duplex MIMO interference channels," *IEEE Trans. Commun.*, vol. 63, no. 3, pp. 801–815, Mar. 2015.

[15] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-Aho, "On the spectral efficiency of full-duplex small cell wireless systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4896–4910, Sep. 2014.

[16] S. Huberman and T. Le-Ngoc, "Full-duplex MIMO precoding for sum-rate maximization with sequential convex programming," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5103–5112, Nov. 2015.

[17] S. Huberman and T. Le-Ngoc, "MIMO full-duplex precoding: A joint beamforming and self-interference cancellation structure," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2205–2217, Apr. 2015.

[18] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-Aho, "Transmission strategies for full duplex multiuser MIMO systems," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Ottawa, ON, Canada, Jun. 2012, pp. 6825–6829.

[19] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-Aho, "Precoding for full duplex multiuser MIMO systems: Spectral and energy efficiency maximization," *IEEE Trans. Signal Process.*, vol. 61, no. 16, pp. 4038–4050, Aug. 2013.

[20] B. Yin, M. Wu, C. Studer, J. R. Cavallaro, and J. Lilleberg, "Full-duplex in large-scale wireless systems," in *Proc. Asilomar Conf. Signals, Syst. Comput.*, Pacific Grove, CA, USA, Nov. 2013, pp. 1623–1627.

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

[22] P. Tsiaflakis, J. Vangorp, J. Verlinden, and M. Moonen, "Multiple access channel optimal spectrum balancing for upstream DSL transmission," *IEEE Commun. Lett.*, vol. 11, no. 4, pp. 300–398, Apr. 2007.

[23] V. L. Nir, M. Moonen, J. Verlinden, and M. Guenach, "Optimal power allocation for downstream xDSL with per-modem total power constraints: Broadcast channel optimal spectrum balancing (BC-OSB)," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 690–697, Feb. 2009.

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

[25] P. Tsiaflakis, R. B. Moraes, and M. Moonen, "A low-complexity algorithm for joint spectrum and signal coordination in upstream DSL transmission," in *Proc. IEEE 18th Symp. Commun. Veh. Technol. Benelux*, Ghent, Belgium, Nov. 2011, pp. 1–6.

[26] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
[27] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.
[28] M. H. M. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inf. Theory*, vol. 29, no. 5, pp. 439–441, May 1983.
[29] W. Yu and R. Lui, "Dual methods for nonconvex spectrum optimization of multicarrier systems," *IEEE Trans. Commun.*, vol. 54, no. 7, pp. 1310–1322, Jul. 2006.
[30] S. Boyd and L. Vandenberghe, "Duality," in *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.
[31] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2646–2660, Jun. 2007.
[32] R. Hunger, M. Joham, and W. Utschick, "On the MSE-duality of the broadcast channel and the multiple access channel," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 698–713, Feb. 2009.
[33] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed. Philadelphia, PA, USA: SIAM, Jan. 2003.
[34] J. Verdyck and M. Moonen, "Dynamic spectrum management in digital subscriber line networks with unequal error protection requirements," *IEEE Access*, vol. 5, pp. 18107–18120, 2017.
[35] W. Lanneer, P. Tsiaflakis, J. Maes, and M. Moonen, "Linear and nonlinear precoding based dynamic spectrum management for downstream vectored G.fast transmission," *IEEE Trans. Commun.*, vol. 65, no. 3, pp. 1247–1259, Mar. 2017.
[36] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM J. Optim.*, vol. 23, no. 2, pp. 1126–1153, Jan. 2013.
[37] 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. Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.
[38] F. Facchinei, G. Scutari, and S. Sagratella, "Parallel selective algorithms for nonconvex big data optimization," *IEEE Trans. Signal Process.*, vol. 63, no. 7, pp. 1874–1889, Apr. 2015.
[39] F. Facchinei, L. Lampariello, and G. Scutari, "Feasible methods for nonconvex nonsmooth problems with applications in green communications," *Math. Program.*, vol. 164, nos. 1–2, pp. 55–90, Jul. 2017.
[40] L. Sorber, M. Van Barel, and L. De Lathauwer, "Unconstrained optimization of real functions in complex variables," *SIAM J. Optim.*, vol. 22, no. 3, pp. 879–898, Jan. 2012.
[41] K. B. Petersen and M. S. Pedersen, *The Matrix Cookbook*. Lyngby, Denmark: Technical Univ. Denmark, Nov. 2012.
[42] S.-J. Kim and G. B. Giannakis, "Optimal resource allocation for MIMO ad hoc cognitive radio networks," *IEEE Trans. Inf. Theory*, vol. 57, no. 5, pp. 3117–3131, May 2011.
[43] G. Scutari, F. Facchinei, and L. Lampariello, "Parallel and distributed methods for constrained nonconvex optimization—Part I: Theory," *IEEE Trans. Signal Process.*, vol. 65, no. 8, pp. 1929–1944, Apr. 2017.
[44] R. Burachik, L. M. G. Drummond, A. N. Iusem, and B. F. Svaiter, "Full convergence of the steepest descent method with inexact line searches," *Optimization*, vol. 32, no. 2, pp. 137–146, Jan. 1995.
[45] P. Tsiaflakis, F. Glineur, and M. Moonen, "Real-time dynamic spectrum management for multi-user multi-carrier communication systems," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 1124–1137, Mar. 2014.
[46] W. Yu, W. Rhee, S. Boyd, and J. M. Cioffi, "Iterative water-filling for Gaussian vector multiple-access channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 1, pp. 145–152, Jan. 2004.
[47] R. Cendrillon, J. Huang, M. Chiang, and M. Moonen, "Autonomous spectrum balancing for digital subscriber lines," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4241–4257, Aug. 2007.
[48] S. Huberman, C. Leung, and T. Le-Ngoc, "Constant offset autonomous spectrum balancing using multiple reference lines for VDSL," *IEEE Trans. Signal Process.*, vol. 60, no. 12, pp. 6719–6723, Dec. 2012.
[49] *Fast Access to Subscriber Terminals (G.Fast)—Power Spectral Density Specification*, document ITU-T G.9700, International Telecommunication Union, ITU-T Study Group 15, 2014.
[50] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, vol. 317, 1st ed. Berlin, Germany: Springer, 2009.



JEROEN VERDYCK (S'15) received the M.Sc. degree in electrical engineering from KU Leuven, Leuven, Belgium, in 2014, where he is currently pursuing the Ph.D. degree with the Electrical Engineering Department, under the supervision of Prof. M. Moonen. He is also involved in joint projects with KU Leuven and the University of Antwerp, Antwerp, Belgium. His research interest includes signal processing and optimization for digital communication systems with an emphasis on DSL wireline access networks.



WOUTER LANNEER received the M.Sc. and Ph.D. degrees in electrical engineering from KU Leuven, Belgium, in 2014 and 2019, respectively. He was a Visiting Researcher with the University of Toronto, in 2018. He is currently a member of Nokia Bell Labs, Antwerp, Belgium. His research interests include MIMO communications, information theory, and signal processing and optimization for digital communication systems with an emphasis on ultra-broadband DSL access networks.



PASCHALIS TSIAFLAKIS received the M.Sc. and Ph.D. degrees in electrical engineering from KU Leuven, Belgium, in 2004 and 2009, respectively. He has further conducted research at Princeton University, UCLA, and the Universite Catholique de Louvain. Since 2013, he has been with Nokia Bell Labs, where he currently serves as a Distinguished Member of Technical Staff. His research interests include on research, contributing to standardization bodies, and driving innovation into next-generation communication systems. He has performed research in the fields of optimization, signal processing, and machine learning, with applications to wireline and wireless communication systems. He received the Belgian Young ICT Personality Award, in 2010, the Nokia Innovation Award, in 2017, the Nokia Bell Top Inventor Award, in 2019, and the Distinguished Member of Technical Staff Award, in 2019.



WERNER COOMANS received the M.Sc. degree in electrical engineering and the Ph.D. degree in engineering sciences from Vrije Universiteit Brussel, in 2009 and 2013, respectively, with a focus on semiconductor lasers. In 2013, he joined Nokia Bell Labs, where he currently heads the Copper Access Department within the Access Laboratory, focusing on next-generation gigabit access technologies leveraging legacy copper infrastructure. He actively contributes to several standardization bodies, such as ITU-T and CableLabs, and was a recipient of a Ph.D. Fellowship of the Research Foundation Flanders (FWO), the Nokia Spotlight Recognition Award, in 2016, and the Best Paper Award at the IEEE GLOBECOM, in 2014.



PANAGIOTIS (PANOS) PATRINOS (M'13) received the M.Eng. degree, the M.S. degree in applied mathematics, and the Ph.D. degree in control and optimization from the National Technical University of Athens, in 2010, 2005, and 2003, respectively.

He held a postdoctoral position at the University of Trento and IMT Lucca, Italy, where he became an Assistant Professor, in 2012. In 2014, he held a visiting assistant professor position with the Department of Electrical Engineering, Stanford University. He is currently an Assistant Professor with the Department of Electrical Engineering (ESAT), KU Leuven, Belgium. His research interests include the interface of optimization, learning, and control. His current research aims at theoretical and algorithmic advances with a particular focus on embedded, large-scale, and distributed nonconvex, nonsmooth optimization and data-driven model predictive control and dynamical decision-making under uncertainty. His algorithms are being applied in a broad range of application areas, including autonomous vehicles, smart grids, water networks, aerospace, multi-agent systems, signal processing, and machine learning.

Dr. Patrinos was the General Chair of the 38th Benelux Meeting on Systems and Control, in 2019, the Co-Chair of the 4th European Conference on Computational Optimization (EUCCO 2016), and a member of the International Program Committee of the 58th IEEE Conference on Decision and Control, in 2019.



MARC MOONEN (M'94–SM'06–F'07) is currently a Full Professor with the Electrical Engineering Department, KU Leuven, where he is heading a research team working in the area of numerical algorithms and signal processing for digital communications, wireless communications, DSL, and audio signal processing.

Dr. Moonen has been a Fellow of the European Association for Signal Processing (EURASIP), since 2018. He received the 1994 KU Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with P. Vandaele), and the 2004 Alcatel Bell (Belgium) Award (with R. Cendrillon) and was a 1997 Laureate of the Belgium Royal Academy of Science. He received the journal best paper awards from the IEEE TRANSACTIONS ON SIGNAL PROCESSING (with G. Leus and with D. Giacobello) and from *Elsevier Signal Processing* (with S. Doclo). He was the Chairman of the IEEE Benelux Signal Processing Chapter, from 1998 to 2002, a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications, and the President of EURASIP, from 2007 to 2008 and from 2011 to 2012. He has served as an Editor-in-Chief for the *EURASIP Journal on Applied Signal Processing*, from 2003 to 2005, and an Area Editor for Feature Articles in the *IEEE Signal Processing Magazine*, from 2012 to 2014, and has been a member of the Editorial Board of *Signal Processing*, the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II, the *IEEE Signal Processing Magazine*, the *Integration-the VLSI Journal*, the *EURASIP Journal on Wireless Communications and Networking*, and the *EURASIP Journal on Advances in Signal Processing*.