IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received June 10, 2019, accepted June 27, 2019, date of publication July 12, 2019, date of current version August 22, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2928451

Physical-Layer Multicasting Design for Downstream G.fast DSL Transmission

WOUTER LANNEER[1], YA-FENG LIU[2], (Senior Member, IEEE), WEI YU[3], (Fellow, IEEE), AND MARC MOONEN[1], (Fellow, IEEE)

[1]ESAT/STADIUS Center for Dynamical Systems, Signal Processing and Data Analytics, KU Leuven, 3000 Leuven, Belgium
[2]State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China
[3]Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada

Corresponding author: Wouter Lanneer (wlanneer@esat.kuleuven.be)

This work was supported in part by the ESAT Laboratory of KU Leuven in the frame of VLAIO O&O Project "the 5th Generation Broadband" under Grant HBC.2016.0055, the project "Multi-gigabit Innovations in Access" under Grant HBC.2017.1007, the Research Project FWO "Real-Time Adaptive Cross-Layer Dynamic Spectrum Management for Fifth Generation Broadband Copper Access Networks" under Grant G.0B1818N, the Fonds de la Recherche Scientifique (FNRS) and the Fonds Wetenschappelijk Onderzoek–Vlaanderen through the EOS Project "MUlti-SErvice WIreless NETwork (MUSE-WINET)" under Grant 30452698. The work of Y.-F. Liu was supported in part by the National Natural Science Foundation of China (NSFC) under Grant 11671419 and Grant 11688101.

ABSTRACT This paper studies the physical-layer multicasting design for downstream G.fast digital subscriber line (DSL) transmission, which corresponds to a multi-user multi-tone (i.e., multi-carrier) scenario. The design goal is to maximize the weighted-sum-group-rate (WSGR) under per-line power constraints. First, as an information-theoretic upper bound, full-rank precoding-based multicasting is considered with joint channel coding across tones. For a single multicast group, this problem corresponds to a non-linear convex semidefinite program (SDP), which is coupled across tones. To reduce the computational complexity, a Lagrange dual decomposition method is developed. This approach is then extended toward multiple multicast groups based on difference-of-convex (DC) programming. Furthermore, a practical multicasting scheme is considered based on rank-one single-stream precoding and independent per-tone channel coding. For this case, instead of relying on computationally complex semidefinite relaxation, a successive convex approximation-based trust-region algorithm is developed. Finally, the simulations of a G.fast cable binder show that the practical multicasting scheme operates close to the information-theoretic multicasting upper bound.

INDEX TERMS G.fast, dynamic spectrum management, physical-layer multicasting, rank-one precoding.

I. INTRODUCTION

Introduced by the International Telecommunication Union (ITU), G.fast [2] is the digital subscriber lines (DSL) access technology that marks the beginning of "ultra-broadband copper access" by offering gigabit (i.e. fiber-like) transmission speeds. These speeds are achieved by employing discrete multi-tone modulation (DMT) in a broad spectrum up to 212 MHz over very short copper telephony lines (below 100 m). Importantly, the use of such high frequencies leads to increasingly stronger levels of crosstalk interference among the lines within a cable binder [3]. This resulted in many advanced precoding-based unicasting strategies for downstream DSL transmission [4]–[6], where all subscribers or users receive independent data streams.

To further improve DSL networks, other transmission strategies that fully take advantage of these strong crosstalk interference have to be developed. To that end, this paper considers physical-layer multicasting, taking into consideration that some users may request the same data streams at the same time. This is for instance the case with IP-multicast based (radio and TV) broadcasting, video conferencing or live event streaming. In such a scenario, physical-layer multicasting is indeed able to outperform standard unicasting in G.fast, due to these strong crosstalk interference, which provides the cable binder with a multiple-input-multiple-output (MIMO) capacity or power gain. A similar scenario appears in the specific case of a cloud radio access network (C-RAN) in small-cell deployment, where the fronthaul or backhaul links

The associate editor coordinating the review of this manuscript and approving it for publication was Bijoy Chand Chatterjee.

between the centralized processor and the base stations are provided by copper telephony lines. Such a C-RAN using the data-sharing cooperation strategy is then able to take advantage of a possible DSL multicasting capability [7], [8]. Moreover, a recent consideration of the DSL community is that future DSL technologies should encompass point-to-multipoint transmission to many in-home access points and devices [9], yielding potential multicasting use cases.

A practical and widely adopted transmission scheme for physical layer multicasting is rank-one precoding, which corresponds to using a single precoder vector and data stream for each multicast group. This scheme has been extensively studied in wireless communications for single-carrier channels. Two basic precoding design problems are sum-power minimization under signal-to-interference-plus-noise ratio (SINR) constraints (the quality-of-service (QoS) problem), and SINR maximization under a sum-power constraint (the max-min-fairness problem). However, even for single-group multicasting, these precoding design problems are non-convex [10]. A first approach to solving such a precoding design problem is based on semidefinite relaxation (SDR), by reformulating the problem as a rank-one constrained semidefinite program (SDP). Hence, subsequently dropping the non-convex rank-one constraint leads to a convex SDP which corresponds in fact to the single-group multicasting capacity [11]. This convex SDP has a possibly full-rank transmit covariance matrix as a solution, which may be approximated by a rank-one matrix using a randomization procedure [12]. This SDR approach has been extended for multi-group multicasting in [13].

On the other hand, directly optimizing the precoding vectors based on successive convex approximation (SCA) [14] has been shown to outperform the SDR approach, both in terms of performance and computational complexity. See e.g. [15] where a locally-optimal iterative second-order cone programming (SOCP) is proposed for the single-group multicasting scenario. Moreover, a globally-optimal branch-and-bound algorithm for single-group multicasting has been proposed in [16]. For multi-group multicasting, on the other hand, a SCA approach has been proposed by relying on the feasible-point-pursuit-SCA algorithm for non-convex quadratically constrained quadratic programs (QCQPs) [17], [18].

Nevertheless, DSL transmission corresponds to a multi-tone (i.e. multi-carrier) scenario, such that in this paper, different from the QoS and max-min-fairness design problems, the weighted-sum-group-rate (WSGR) maximization design problem for multi-group multicasting is studied. This problem has been tackled in [19], [20] by a two-step heuristic algorithm for a single-carrier scenario. However, the first step corresponds to the QoS problem given fixed SINR constraints which is being solved by the computationally complex SDR approach. The second step consists of a power reallocation given fixed precoding vectors via a subgradient approach. In addition, a SCA approach has been recently proposed for maximizing the minimum group rate with antenna selection for multi-carrier systems with (a limited number) of sub-carriers in [21]. However, the large number of tones in DSL together with the design problem being coupled across the tones leads to excessive computational complexity, which is specifically addressed in this paper.

A. MAIN CONTRIBUTIONS

In this paper, the physical-layer multicasting design problem of maximizing the WSGR is considered for downstream G.fast DSL transmission, which corresponds to a multi-user multi-tone (i.e. multi-carrier) scenario.

The first part of this paper considers full-rank precoding-based multicasting, by means of full-rank transmit covariance matrices in combination with joint channel coding across tones. This means that users of the same multicast group communicate at the same total bit-rate aggregated over all tones, while they may have a different set of SINR values across the tones. For the single-group case, this leads to a non-linear convex SDP which is coupled across tones, corresponding to the multicasting capacity. To deal with the large number of tones in G.fast, a Lagrange dual decomposition method with a subgradient search is proposed for solving this SDP. This approach is then generalized for multi-group multicasting by relying on SCA-SDP, based on difference-of-convex (DC) programming. Unfortunately, since a practical implementation of this multicasting scheme has not yet been realized, it is information-theoretic in nature and merely serves as an upper bound for practical multicasting schemes.

The second part of this paper considers a practical rank-one single-stream precoding based multicasting scheme, together with independent per-tone channel coding, such that users of the same multicast group communicate at the same bit-rate on every tone independently, according to the minimum SINR value. For this case a trust-region method based on SCA is proposed to maximize the WSGR. In addition, inspired by [22], a zero-forcing (ZF) rank-one precoding based multicasting scheme is proposed based on the two-layer block-diagonalizing precoder of [23], [24], to further reduce the computational complexity.

Finally, simulations of downstream transmission in a 10-line G.fast cable binder are provided. The practical multicasting scheme is shown to operate close to the information-theoretic upper bound.

B. ORGANIZATION AND NOTATION

This paper is organized as follows. Section II introduces the system model for downstream G.fast DSL transmission. Section III addresses WSGR maximization for full-rank precoding-based multicasting, while Section IV considers rank-one precoding-based multicasting. Section V presents simulation results for a G.fast cable binder. Finally, Section VI concludes the paper.

Lower-case boldface letters are used to denote vectors and uppercase boldface letters for matrices. Further, $\mathbf{I}_A$ is used as the identity matrix of size A, $(.)^T$ as the transpose, $(.)^{\mathcal{H}}$ as the Hermitian transpose, $(.)^*$ as the complex conjugate, $\mathcal{E}\{.\}$ as

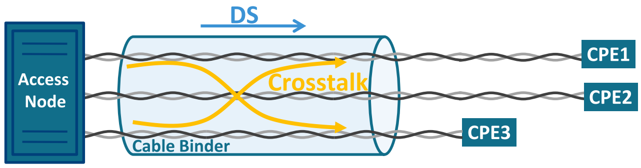


FIGURE 1. A multi-user G.fast DSL network typically experiences strong crosstalk interference at high frequencies between the copper lines.

expectation,$[\mathbf{X}]_{ij}$ as the i, j-th element of $\mathbf{X}$, $\mathrm{Tr}\{.\}$ as trace, $\mathrm{diag}(\mathbf{x})$ a diagonal matrix with $\mathbf{x}$ on the diagonal, $|\cdot|$ as the scalar absolute value, and $\mathbf{1}$ as a vector of ones.

II. SYSTEM MODEL

Downstream transmission in a G.fast cable binder with N lines or users is considered (see Fig. 1). Assuming standard synchronous DMT with K tones and a sufficiently long cyclic prefix, the transmission is modeled independently across the tones as

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{z}_k, \quad \text{for } k = [1, \ldots, K], \tag{1}$$

where $\mathbf{x}_k \triangleq [x_k^1, \ldots, x_k^N]$ is the transmit vector on tone k, with x_k^n the signal transmitted on line n, $\mathbf{y}_k \triangleq [y_k^1, \ldots, y_k^N]$ is the receive vector on tone k, with y_k^n the signal received by user n, $\mathbf{z}_k \triangleq [z_k^1, \ldots, z_k^N]$ is the vector of uncorrelated additive noise signals on tone k, with unity noise power, i.e., $\mathcal{E}\{|z_k^n|^2\} \triangleq 1$. The $N \times N$ complex channel matrix on tone k is denoted by

$$\mathbf{H}_k \triangleq \begin{bmatrix} \mathbf{h}_k^{1,\mathcal{H}} \\ \vdots \\ \mathbf{h}_k^{N,\mathcal{H}} \end{bmatrix}, \tag{2}$$

where the $\mathbf{h}_k^{n,\mathcal{H}}$ is the channel row vector from the access node to the receiver of user n. The diagonal elements of $\mathbf{H}_k$ represent the direct channels, whereas the off-diagonal elements represent the crosstalk channels. Although the direct channels of $\mathbf{H}_k$ typically are dominant below 30 MHz (i.e. $|[\mathbf{H}_k]_{nn}| \gg |[\mathbf{H}_k]_{nm}|, m \neq n$), this is not valid for higher frequencies of G.fast where the direct channels may even be weaker than the crosstalk channels [3].

Perfect knowledge of the channel matrices is assumed. In DSL systems the channel characteristics vary slowly with time such that the access node is indeed able to estimate and track the channel characteristics by sending pilot symbols interleaved with the data symbols [2].

In G.fast, per-line spectral mask constraints and aggregate transmit power (ATP) constraints are enforced, i.e.,

$$\mathcal{E}\{|x_k^n|^2\} \leq P_k^{\text{mask}}, \quad \forall k, n \tag{3}$$

$$\sum_k \mathcal{E}\{|x_k^n|^2\} \leq P^{\text{ATP}}, \quad \forall n. \tag{4}$$

The capacity of this multi-tone channel for user n in bits/s is then given by $R_n = f_s \sum_k b_k^n$ where f_s is the DMT symbol rate and b_k^n represents the achievable bit-rate (in bits per DMT symbol) on tone k

$$b_k^n = \log_2\left(1 + \mathrm{SINR}_k^n\right) \tag{5}$$

with SINR_k^n the SINR at the receiver of user n on tone k. Note that in Section III and IV the natural logarithm is adopted in (5) and f_s is dropped for concise notation, such that all bit-rates in those sections are expressed in nats/DMT symbol.

III. FULL-RANK PRECODING-BASED MULTICASTING

In this section, a dual decomposition algorithm for full-rank precoding-based multicasting is proposed, which corresponds to the multicasting capacity for a single group. Second, the proposed algorithm is generalized to multi-group multicasting, by relying on DC programming.

A. SINGLE-GROUP MULTICASTING

From an information-theoretic perspective, the single-group multicasting capacity for model (1) under per-line power constraints can be shown to be [11], [25]

$$\begin{aligned} \underset{\{\mathbf{C}_k \succeq \mathbf{0}\}}{\text{maximize}} \quad & \min_n \left\{ \sum_k \log\left(1 + \mathbf{h}_k^{n,\mathcal{H}} \mathbf{C}_k \mathbf{h}_k^n\right) \right\} \\ \text{s.t.} \quad & [\mathbf{C}_k]_{nn} \leq P_k^{\text{mask}}, \quad \forall k, n \\ & \sum_k [\mathbf{C}_k]_{nn} \leq P^{\text{ATP}}, \quad \forall n \end{aligned} \tag{6}$$

where $\mathbf{C}_k \triangleq \mathcal{E}\{\mathbf{x}_k \mathbf{x}_k^{\mathcal{H}}\}$ is the $N \times N$ positive-semidefinite transmit covariance matrix on tone k. Observe that (6) assumes joint channel coding across tones. All users thus communicate at the same minimum bit-rate aggregated over all tones, while they may achieve a different set of SINR values across the tones. The use of independent channel coding on each tone separately, on the other hand, restricts users from the same multicast group to communicate at the same bit-rate at every tone separately, according to the minimum SINR value. This is equivalent to shifting the summation over the tones k outside the min-function in the objective of (6). Although the use of independent channel coding is in fact optimal with respect to the unicasting capacity[1] [26, Ch. 5.3.3], this is not valid for the case of multicasting.

Note that problem (6) does not assume any physical-layer multicasting scheme, it merely assumes possibly full-rank transmit covariance matrices. As a result, the capacity promised by (6) can be considered as an upper bound for any possible physical-layer multicasting scheme. However, for a single tone ($K = 1$) channel with up to $N = 3$ users, rank-one single-stream precoding is capacity-achieving since the optimal transmit covariance matrix in (6) is rank-one [11], [27]. In addition, rank-2 Alamouti precoding is capacity-achieving up to $N = 8$ users on a single tone [11].

[1] However, for practical (non-capacity achieving) codes, channel coding across the tones may improve the bit error probability for a certain bit-rate.

Since the minimum of a set of concave functions is concave, (6) is a (non-smooth) convex problem. Notwithstanding its convexity, (6) has a high complexity, due to the coupling across tones and the large number of tones K in DSL networks.[2] Therefore, (6) is first reformulated into a smooth problem and then decoupled into K independent low-complexity subproblems, by relying on Lagrange dual decomposition.

The smooth version of (6) is formulated by introducing the auxiliary (positive real) variable t

$$\underset{t,\{\mathbf{C}_k\succeq\mathbf{0}\}}{\text{maximize}} \quad t \tag{7a}$$

$$\text{s.t.} \ \sum_k \log\left(1+\mathbf{h}_k^{n,\mathcal{H}}\mathbf{C}_k\mathbf{h}_k^n\right) \geq t \quad \forall n \tag{7b}$$

$$[\mathbf{C}_k]_{nn} \leq P_k^{\text{mask}}, \quad \forall k,n \tag{7c}$$

$$\sum_k [\mathbf{C}_k]_{nn} \leq P^{\text{ATP}}, \quad \forall n. \tag{7d}$$

Since (7) is convex with strictly feasible constraints, strong duality holds in (7) and Lagrange dual decomposition may be used to decouple the constraint functions in (7b) and (7d). Towards this end, define the non-negative Langrange multipliers $\boldsymbol{\theta} \triangleq [\theta_1,\ldots,\theta_N]$ and $\boldsymbol{\lambda} \triangleq [\lambda_1,\ldots,\lambda_N]$ corresponding to (7b) and (7d), respectively. Then, the Lagrangian is formed by augmenting the objective in (7a) with the weighted sum of the constraint functions [29, Ch. 5]

$$\begin{aligned}\mathcal{L}&(t,\{\mathbf{C}_k\},\boldsymbol{\theta},\boldsymbol{\lambda})\\ &= t-\sum_n \theta_n\left(t-\sum_k \log\left(1+\mathbf{h}_k^{n,\mathcal{H}}\mathbf{C}_k\mathbf{h}_k^n\right)\right)\\ &\quad -\sum_n \lambda_n\left(\sum_k [\mathbf{C}_k]_{nn}-P^{\text{total}}\right)\end{aligned} \tag{8}$$

$$\begin{aligned}&=\left(1-\sum_n\theta_n\right)t+\sum_n\sum_k \theta_n \log\left(1+\mathbf{h}_k^{n,\mathcal{H}}\mathbf{C}_k\mathbf{h}_k^n\right)\\ &\quad -\sum_n \lambda_n\left(\sum_k [\mathbf{C}_k]_{nn}-P^{\text{total}}\right)\end{aligned} \tag{9}$$

and the Lagrange dual problem is

$$\underset{\boldsymbol{\theta}\succeq\mathbf{0},\boldsymbol{\lambda}\succeq\mathbf{0}}{\text{minimize}}\ \underset{t,\{\mathbf{C}_k\succeq\mathbf{0}\}}{\text{maximize}}\ \mathcal{L}(t,\{\mathbf{C}_k\},\boldsymbol{\theta},\boldsymbol{\lambda}). \tag{10}$$

If $(1-\sum_n\theta_n)$ is positive, it is easy to see that the Lagrangian is unbounded from above (i.e. setting $t=\infty$ is optimal), meaning that the Lagrange dual problem (10) is infeasible. If, on the other hand, $(1-\sum_n\theta_n)$ is negative, the optimal t has zero value, meaning that the primal problem (7) is infeasible. This results in the hidden constraint $\sum_n\theta_n = 1$ and yields the following equivalent Lagrange dual problem (where t has been effectively removed):

$$\begin{aligned}&\underset{\boldsymbol{\theta}\succeq\mathbf{0},\boldsymbol{\lambda}\succeq\mathbf{0}}{\text{minimize}}\ \sum_k g_k(\boldsymbol{\theta},\boldsymbol{\lambda})+\boldsymbol{\lambda}^T\mathbf{P}^{\text{total}}\\ &\text{s.t.}\ \sum_n \theta_n = 1.\end{aligned} \tag{11}$$

where $\mathbf{P}^{\text{ATP}} \triangleq \mathbf{1}P^{\text{ATP}}$. The Lagrange dual function consists of K independent per-tone subproblems $g_k(\boldsymbol{\theta},\boldsymbol{\lambda})$

$$\begin{aligned}&g_k(\boldsymbol{\theta},\boldsymbol{\lambda})\\ &=\underset{\mathbf{C}_k\succeq\mathbf{0}}{\text{maximize}}\ \sum_n \theta_n\log\left(1+\mathbf{h}_k^{n,\mathcal{H}}\mathbf{C}_k\mathbf{h}_k^n\right)-\text{Tr}\left\{\text{diag}(\boldsymbol{\lambda})\mathbf{C}_k\right\}\\ &\quad\text{s.t.}\ [\mathbf{C}_k]_{nn}\leq P_k^{\text{mask}}, \quad \forall n.\end{aligned} \tag{12}$$

Each per-tone subproblem (12) is a smooth small-scale convex problem and thus may be efficiently solved to its global optimum with a standard optimization tool like e.g. CVX [30].

The remaining difficulty lies in finding the optimal Lagrange multipliers $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$ that solve the equivalent dual problem (11). A well-known technique for solving this dual problem is the projected subgradient method [31]. A possible subgradient direction for $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$ at iterate i is

$$\mathbf{d}_{\boldsymbol{\theta}}^{(i)} = \mathbf{R}^{(i)} \tag{13}$$

$$\mathbf{d}_{\boldsymbol{\lambda}}^{(i)} = \mathbf{P}^{\text{ATP}}-\sum_k \text{diag}\left(\mathbf{C}_k^{(i)}\right) \tag{14}$$

where $\mathbf{C}_k^{(i)}$ is the optimized transmit covariance matrix in $g_k\left(\boldsymbol{\theta}^{(i)},\boldsymbol{\lambda}^{(i)}\right)$, which yields $\mathbf{R}^{(i)} \triangleq [R_1^{(i)},\ldots,R_N^{(i)}]^T$ with $R_n^{(i)} = \sum_k \log\left(1+\mathbf{h}_k^{n,\mathcal{H}}\mathbf{C}_k^{(i)}\mathbf{h}_k^n\right)$. The basic update of the Lagrange multipliers in the subgradient direction with step sizes $\delta_{\boldsymbol{\theta}}^i$ and $\delta_{\boldsymbol{\lambda}}^i$ is then written as

$$\boldsymbol{\theta}^{(i+1)} = \boldsymbol{\theta}^{(i)}-\delta_{\boldsymbol{\theta}}^i\mathbf{d}_{\boldsymbol{\theta}}^{(i)} \tag{15}$$

$$\boldsymbol{\lambda}^{(i+1)} = \boldsymbol{\lambda}^{(i)}-\delta_{\boldsymbol{\lambda}}^i\mathbf{d}_{\boldsymbol{\lambda}}^{(i)}. \tag{16}$$

The updated Lagrange multipliers should be Euclidean projected back on their respective constraint sets. For $\boldsymbol{\lambda}$, this projection is realized by simply replacing (16) with

$$\boldsymbol{\lambda}^{(i+1)} = \left[\boldsymbol{\lambda}^{(i)}-\delta_{\boldsymbol{\lambda}}^i\mathbf{d}_{\boldsymbol{\lambda}}^{(i)}\right]^+. \tag{17}$$

For $\boldsymbol{\theta}$, on the other hand, there is a joint Euclidean projection needed onto two sets: $C^1 = \{\boldsymbol{\theta}|\mathbf{1}^T\boldsymbol{\theta}=1\}$ and $C^2 = \{\boldsymbol{\theta}|\boldsymbol{\theta}\succeq\mathbf{0}\}$, which is detailed in Appendix A. Note that the Euclidean projection onto the constraint set never results in moving further away from the optimal point. For instance, if all obtained data rates corresponding to $\boldsymbol{\theta}^{(i)}$ are equal at a certain iteration i (i.e., $\boldsymbol{\theta}^{(i)}$ is the optimal Lagrange multiplier vector with $\mathbf{R}^{(i)} = R\mathbf{1}_N$), then Euclidean projection of $\boldsymbol{\theta}^{(i+1)}$ results into the same vector $\boldsymbol{\theta}^{(i)}$. Further, the projected subgradient method is guaranteed to converge if the step sizes $\delta_{\boldsymbol{\theta}}^i$ and $\delta_{\boldsymbol{\lambda}}^i$ are chosen sufficiently small [31].

The complete algorithm is summarized in Alg. 1 and is guaranteed to converge to the global optimum of the convex

[2]Note that as the variable size scales with K in (6), the worst-case complexity per iteration with interior-point methods scales with $\mathcal{O}(K^3)$ [28], where the maximum number of tones in G.fast is $K = 4096$.

Algorithm 1 Single-Group Multicasting

Initialize $\boldsymbol{\theta} = \mathbf{1}/N$ and $\boldsymbol{\lambda} \succeq \mathbf{0}$
repeat
  Update $\boldsymbol{\theta} = \boldsymbol{\theta} - \delta_{\boldsymbol{\theta}} \mathbf{d}_{\boldsymbol{\theta}}$
  Euclidean project $\boldsymbol{\theta}$ onto the constraint set (see Appendix A)
  repeat
    Update $\boldsymbol{\lambda} = [\boldsymbol{\lambda} - \delta_{\boldsymbol{\lambda}} \mathbf{d}_{\boldsymbol{\lambda}}]^+$
    for $k = 1 \cdots K$ **do**
      Update $\mathbf{C}_k$ by solving $g_k(\boldsymbol{\theta}, \boldsymbol{\lambda})$ in (12)
    end
  until $\lambda^n \left|\sum_k [\mathbf{C}_k]_{nn} - P^{\text{total}}\right| < \epsilon \quad \forall n$
  Update $R_n^* = \sum_k \log\left(1 + \mathbf{h}_k^{n,\mathcal{H}} \mathbf{C}_k \mathbf{h}_k^n\right) \quad \forall n$
until $\theta^n \left|\min_m\{R_m^*\} - R_n^*\right| < \epsilon \quad \forall n$

single-group multicasting problem (6). For a general number of users N and tones K, the rank of the optimal $\mathbf{C}_k$ in (6) may be larger than one. However, in case of a single user ($N = 1$), the single optimal Lagrange multiplier θ^* equals one and the per-tone slave problems $g_k(\theta^*, \boldsymbol{\lambda})$ have an optimal rank-one solution (see e.g. [32] for an SVD-based method). Furthermore, in case of a single tone ($K = 1$), problem (7) may be simplified to a linear SDP [10].

B. MULTI-GROUP MULTICASTING

In this subsection, Alg. 1 is generalized for single-group multicasting towards the case with multiple interfering multicasting groups. Consider G groups with $1 \le G \le N$, and let $\mathcal{G}_g$ denote the set of users in group g, and $|\mathcal{G}_g|$ the number of users in group g. Each user n is member of one and only one group, denoted by $g_n \in [1, \ldots, G]$. Define $\mathbf{C}_k^g \triangleq \mathcal{E}\{\mathbf{x}_k^g \mathbf{x}_k^{g,\mathcal{H}}\}$ as the $N \times N$ transmit covariance matrix of group g on tone k, with $\mathbf{x}_k = \sum_g \mathbf{x}_k^g$ and $\mathbf{C}_k = [\mathbf{C}_k^1, \ldots, \mathbf{C}_k^G]$. Then the achievable bit-rate of user n on tone k is

$$b_k^n(\mathbf{C}_k) \triangleq \log\left(1 + \frac{\mathbf{h}_k^{n,\mathcal{H}} \mathbf{C}_k^{g_n} \mathbf{h}_k^n}{1 + \sum_{j \neq g_n} \mathbf{h}_k^{n,\mathcal{H}} \mathbf{C}_k^j \mathbf{h}_k^n}\right). \tag{18}$$

The goal is to maximize the WSGR by optimizing the transmit covariance matrices under per-line power constraints:

$$\underset{\{\mathbf{C}_k^g \succeq \mathbf{0}\}}{\text{maximize}} \sum_g \alpha_g \min_{n \in \mathcal{G}_g} \left\{\sum_k b_k^n(\mathbf{C}_k)\right\} \tag{19a}$$

$$\text{s.t.} \quad \sum_g [\mathbf{C}_k^g]_{nn} \le P_k^{\text{mask}}, \quad \forall n, k, \tag{19b}$$

$$\sum_k \sum_g [\mathbf{C}_k^g]_{nn} \le P^{\text{total}}, \quad \forall n, \tag{19c}$$

where α_g is the nonnegative weight of group g. Strictly speaking, (19) is not equivalent to the true multi-group multicasting capacity, which requires dirty-paper coding [33] among the groups. However, (19) can be seen as a theoretical achievable rate for any physical-layer scheme using only linear encoding and decoding of the groups. From an optimization point of view, (19) can also be seen as the SDR upper bound of the rank-one single stream precoding scheme in Section IV.

Unlike the single-group case, (19) is non-convex due the inter-group interference. Fortunately, the difference of concave functions structure in (18) allows to leverage successive lower bound maximization (also known as the majorization-minimization (MM) method). To see this, define $\mathbf{A}_k^n \triangleq \mathbf{h}_k^n \mathbf{h}_k^{n,\mathcal{H}}$ and the noise and interference covariance matrix $X_k^n \triangleq 1 + \sum_{j \neq g_n} \text{Tr}\left\{\mathbf{A}_k^n \mathbf{C}_k^j\right\}$, such that (18) may be re-written as follows:

$$b_k^n(\mathbf{C}_k) = \log\left(X_k^n + \text{Tr}\left\{\mathbf{A}_k^n \mathbf{C}_k^{g_n}\right\}\right) - \log\left(X_k^n\right). \tag{20}$$

A first-order approximation of the second term in (20) in $\mathbf{C}_k^g$ around $\overline{\mathbf{C}}_k$, i.e.,

$$\log\left(X_k^n\right) \le \underbrace{\log\left(\overline{X}_k^n\right) - \frac{\text{Tr}\left\{\mathbf{A}_k^n \overline{\mathbf{C}}_k^{g_n}\right\}}{\overline{X}_k^n}}_{C_k^n} + \frac{\text{Tr}\left\{\mathbf{A}_k^n \mathbf{C}_k^{g_n}\right\}}{\overline{X}_k^n}, \tag{21}$$

leads to a global concave lower bound for $b_k^n(\mathbf{C}_k)$

$$\tilde{b}_k^n\left(\mathbf{C}_k | \overline{\mathbf{C}}_k\right) = \log\left(X_k^n + \text{Tr}\left\{\mathbf{A}_k^n \mathbf{C}_k^{g_n}\right\}\right) - \frac{\text{Tr}\left\{\mathbf{A}_k^n \mathbf{C}_k^{g_n}\right\}}{\overline{X}_k^n} - C_k^n. \tag{22}$$

Then, in iteration l, the successive lower bound maximization updates the transmit covariance matrices, by solving the following non-smooth convex problem given the iterate $\left\{\mathbf{C}_k^{(l-1)} | \forall k\right\}$ from the previous iteration:

$$\left\{\mathbf{C}_k^{(l)} | \forall k\right\} = \underset{\{\mathbf{C}_k^g \succeq \mathbf{0}\}}{\arg\max} \sum_g \alpha_g \min_{n \in \mathcal{G}_g} \left\{\sum_k \tilde{b}_k^n\left(\mathbf{C}_k | \mathbf{C}_k^{(l-1)}\right)\right\}$$
$$\text{s.t. (19b) and (19c).} \tag{23a}$$

Problem (23) may be again optimally solved using the same approach as in Section III-A for the single-group case. First the smooth version of (23) is formulated by introducing the auxiliary variables $\{t_g\}$ for each group, i.e.,

$$\underset{\{t_g\}, \{\mathbf{C}_k^g \succeq \mathbf{0}\}}{\text{maximize}} \sum_g \alpha_g t_g$$
$$\text{s.t.} \quad \sum_k \tilde{b}_k^n\left(\mathbf{C}_k | \mathbf{C}_k^{(l-1)}\right) \ge t_g, \quad \forall g, n \in \mathcal{G}_g$$
$$\text{(19b) and (19c).} \tag{24}$$

Since strong duality holds in (24), it can be shown to be equivalent to the following constrained Lagrange dual problem:

$$\underset{\{\boldsymbol{\theta}_g \succeq \mathbf{0}\}, \boldsymbol{\lambda} \succeq \mathbf{0}}{\text{minimize}} \sum_k g_k(\{\boldsymbol{\theta}_g\}, \boldsymbol{\lambda}) + \boldsymbol{\lambda}^T \mathbf{P}^{\text{ATP}}$$
$$\text{s.t.} \sum_{m=1}^{|\mathcal{G}_g|} \theta_{(g,m)} = \alpha_g, \quad \forall g. \tag{25}$$

where $\boldsymbol{\theta}_g \triangleq [\theta_{(g,1)}, \ldots, \theta_{(g,|\mathcal{G}_g|)}]^T$ are the Lagrange multipliers of the $|\mathcal{G}_g|$ users in group g. Moreover, θ_n denotes the Lagrange multiplier associated with user n of group g_n.

Algorithm 2 Multi-Group Multicasting

Initialize $\boldsymbol{\theta}_g = \alpha_g \mathbf{1}/|\mathcal{G}_g|, \forall g$ and $\boldsymbol{\lambda} \succeq \mathbf{0}$
Set $l = 0$ and initialize $\{\mathbf{C}_k^{(l)}\}$ with feasible values
repeat
  Set $l = l + 1$
  repeat
    Update $\boldsymbol{\theta}_g = \boldsymbol{\theta}_g - \delta_{\boldsymbol{\theta}_g} \mathbf{d}_{\boldsymbol{\theta}_g}, \forall g$
    Euclidean project all $\boldsymbol{\theta}_g$ onto the contraint set (App. A)
    repeat
      Update $\boldsymbol{\lambda} = [\boldsymbol{\lambda} - \delta_{\boldsymbol{\lambda}} \mathbf{d}_{\boldsymbol{\lambda}}]^+$
      for $k = 1 \cdots K$ **do**
        Update $\mathbf{C}_k^{(l)}$ by solving $g_k(\{\boldsymbol{\theta}_g\}, \boldsymbol{\lambda})$ in (28)
      end
    until $\lambda^n \left|\sum_k \sum_g \left[\mathbf{C}_k^{g,(l)}\right]_{nn} - P^{\text{ATP}}\right| < \epsilon, \quad \forall n$
    Update
    $R_n^* = \sum_k \tilde{b}_k^n \left(\mathbf{C}_k^{(l)} | \mathbf{C}_k^{(l-1)}\right), \quad \forall g, n \in \mathcal{G}_g$
  until $\theta_n \left|\min_{m \in \mathcal{G}_{g_n}} \{R_m^*\} - R_n^*\right| < \epsilon, \quad \forall n$
until *the objective value* (19a) *converges*

The corresponding subgradient direction and basic update of $\boldsymbol{\theta}_g$ for group g in iteration i are given by

$$\mathbf{d}_{\boldsymbol{\theta}_g}^{(i)} = \mathbf{R}_g^{(i)} = [R_{(g,1)}^{(i)}, \ldots, R_{(g,|\mathcal{G}_g|)}^{(i)}]^T \tag{26}$$

$$\boldsymbol{\theta}_g^{(i+1)} = \boldsymbol{\theta}_g^{(i)} - \delta_{\boldsymbol{\theta}_g}^i \mathbf{d}_{\boldsymbol{\theta}_g}^{(i)} \tag{27}$$

with $R_{(g,m)}^{(i)} \triangleq R_n^{(i)} = \sum_k \tilde{b}_k^n \left(\mathbf{C}_k^{(i)} | \mathbf{C}_k^{(l-1)}\right)$ where $\mathbf{C}_k^{(i)}$ are the optimized transmit covariance matrices, obtained by solving K independent per-tone smooth convex subproblems:

$$\begin{aligned} & g_k(\{\boldsymbol{\theta}_g^{(i)}\}, \boldsymbol{\lambda}^{(i)}) \\ &= \underset{\{\mathbf{C}_k^g \succeq \mathbf{0}\}}{\text{maximize}} \sum_g \left\{ \sum_{n \in \mathcal{G}_n} \theta_n^{(i)} \tilde{b}_k^n (\mathbf{C}_k | \mathbf{C}_k^{(l-1)}) - \text{Tr}\left\{\text{diag}(\boldsymbol{\lambda}^{(i)}) \mathbf{C}_k^g\right\} \right\} \\ & \text{s.t. } [\mathbf{C}_k]_{nn} \le P_k^{\text{mask}}, \quad \forall n. \end{aligned} \tag{28}$$

The complete algorithm for multi-group multicasting is summarized in Alg. 2. The following convergence result for Alg. 2 is stated below:

Theorem 1: The sequence of iterates $\{\mathbf{C}^{(l)}\}$ with $\mathbf{C}^{(l)} \triangleq \{\mathbf{C}_k^{(l)} | \forall k\}$ has a monotonically non-decreasing objective value in (19a), and moreover, the iterates $\{\mathbf{C}^{(l)}\}$ are guaranteed to converge to the set of stationary points of (19).

Proof: Alg. 2 is a special case of the non-smooth successive upper bound minimization algorithm discussed in [34]. For completeness, a full proof is provided in Appendix B. □

With respect to computational complexity, updating the transmit covariance matrices in each iteration by solving problem (28) for all tones is the most expensive step of Alg. 2. It corresponds to a non-linear SDP with G matrices of size $N \times N$ for every tone k, which may be solved by interior point methods of standard solvers such as CVX [30]. Worst-case, interior point methods require $\mathcal{O}(\sqrt{N^2 G} \log(1/\epsilon))$ iterations to solve (28) for each tone up to accuracy ϵ, with for each iteration an approximate computational complexity of $\mathcal{O}((N^2 G)^3)$ [28]. The total complexity of Alg. 2 is hence $\mathcal{O}(I_1\, I_2\, K \sqrt{N^2 G} N^6 G^3 \log(1/\epsilon))$, with I_1 the number of SCA outer iterations, and I_2 the number of Lagrange multiplier update iterations.

Remark 1: Alg. 2 may be seen as the multicasting generalization of the primal domain DSB algorithm[3] presented in [35], dedicated to weighted-sum-rate maximization for standard downstream unicasting [corresponding to a so-called broadcast channel (BC)]. That is, in case of N single-user groups, the optimal Lagrange multiplier vector $\boldsymbol{\theta}^*$ in (25) has a closed-form solution, given by $\boldsymbol{\theta}^* = [\alpha_1, \ldots, \alpha_N]^T$, such that the K independent subproblems (28) correspond to a weighted-sum-rate maximization in a downstream unicasting scenario. Remarkably, for the unicasting case, these subproblems (28) admit optimal rank-one solutions [35], which is not true in general for the multicasting case.

IV. RANK-ONE PRECODING-BASED MULTICASTING

The bit-rates promised by full-rank precoding-based multicasting, as studied in Section III, provide a theoretical achievability, but to the best our knowledge, there is no physical-layer multicasting scheme successfully implemented and demonstrated that is able to practically achieve these bit-rates for any number of users N and tones K. In contrast, a widely adopted practical multicasting scheme is based on rank-one precoding (i.e. using a single precoder vector and a single stream for each multicast group on every tone), which is efficiently implementable, but achieves only sub-optimal bit-rates in general.

In the rank-one precoding case, the transmit vector on tone k is

$$\mathbf{x}_k = \mathbf{P}_k \mathbf{u}_k \tag{29}$$

where $\mathbf{P}_k \triangleq [\mathbf{p}_k^1, \ldots, \mathbf{p}_k^G]$ is the $N \times G$ (complex) precoder matrix, with $1 \le G \le N$ multicast groups. $\mathbf{u}_k \triangleq [u_k^1, \ldots, u_k^G]$ is the data symbol vector, where u_k^g denotes the data symbol of group g on tone k. The data vector is assumed to be independently and identically distributed (i.i.d.) with normalized powers, i.e., $\mathcal{E}\{\mathbf{u}_k \mathbf{u}_k^{\mathcal{H}}\} = \mathbf{I}_G$. The achievable bit-rate of user n on tone k is

$$b_k^n(\mathbf{P}_k) \triangleq \log\left(1 + \frac{|\mathbf{h}_k^{n,\mathcal{H}} \mathbf{p}_k^{g_n}|^2}{1 + \sum_{j \neq g_n} |\mathbf{h}_k^{n,\mathcal{H}} \mathbf{p}_k^j|^2}\right). \tag{30}$$

Another practical assumption adopted in this section is independent per-tone channel coding. Users of the same group hence communicate at the same bit-rate on every

[3] Although the primal domain DSB algorithm in [35, Alg. 3] applies to a full-duplex scenario, for the downstream scenario a specific instance of the algorithm is obtained by setting all user weights in the upstream direction equal to zero.

tone separately, according to the minimum SINR on every tone. Note that this ensures compatibility with practical DSL systems in which QAM symbols are transmitted instead of Gaussian signals, which may be modeled by inserting a SNR gap approximation in (30).[4] In addition, the per-line ATP constraints are dropped, since these are always observed to be inactive in our simulations (Section V). As a result, the corresponding WSGR maximization problem fully decouples into independent per-tone subproblems, which reduces both the implementation and the optimization complexity. However, it is stressed that both per-line ATP and joint channel coding across tones may be included using the same Lagrange dual decomposition approach as in Section III-B.

In the remainder of this section, an iterative trust-region method for WSGR maximization is proposed, based on SCA instead of computationally complex SDR. Then, to further reduce the computational complexity, ZF rank-one precoding is considered.

A. GENERAL RANK-ONE MULTICASTING

The WSGR maximization problem subject to per-line power constraints is formulated independently for every tone k as follows:

$$\underset{\mathbf{P}_k}{\text{maximize}} \sum_g \alpha_g c_k^g \tag{31a}$$

$$\text{s.t. } c_k^g = \min_{n \in \mathcal{G}_g} \left\{ b_k^n\left(\mathbf{P}_k\right) \right\}, \quad \forall g, \tag{31b}$$

$$\left[\mathbf{P}_k \mathbf{P}_k^{\mathcal{H}}\right]_{nn} \leq P_k^{\text{mask}}, \quad \forall n, \tag{31c}$$

where α_g is the weight of group g, and c_k^g is the bit-rate for group g on tone k.

To solve (31), it is first re-formulated into a smooth problem by introducing a set of auxiliary (positive real) variables $\{t_k^g\}$:

$$\underset{t_k^g, \mathbf{P}_k}{\text{maximize}} \sum_g \alpha_g t_k^g \tag{32a}$$

$$\text{s.t. } b_k^n\left(\mathbf{P}_k\right) \geq t_k^g, \quad \forall k, g, n \in \mathcal{G}_g \tag{32b}$$

$$\left[\mathbf{P}_k \mathbf{P}_k^{\mathcal{H}}\right]_{nn} \leq P_k^{\text{mask}}, \quad \forall n, \tag{32c}$$

Solving problem (32) is made difficult by the non-convex constraint (32b) that needs to be tackled. First, the non-convex term $b_k^n(\mathbf{P}_k)$ in (30) is re-written as

$$b_k^n\left(\mathbf{P}_k\right) = \log\left(X_k^n + S_k^n\right) - \log\left(X_k^n\right) \tag{33}$$

with

$$S_k^n = \mathbf{p}_k^{g_n, \mathcal{H}} \mathbf{A}_k^n \mathbf{p}_k^{g_n} \tag{34a}$$

$$X_k^n = 1 + \sum_{j \neq g_n} \mathbf{p}_k^{j, \mathcal{H}} \mathbf{A}_k^n \mathbf{p}_k^j \tag{34b}$$

[4]Although in DSL the coded bits of a codeword are typically mapped into QAM symbols that are transmitted over different tones (i.e. channel coding across tones), the error rate performance is very close to and thus accurately approximated by the per-tone channel coding case [36].

and $\mathbf{A}_k^n = \mathbf{h}_k^n \mathbf{h}_k^{n,\mathcal{H}}$. Unfortunately, since both log-terms in (33) are non-convex due to the quadratic terms in (34a), there is no longer a difference of concave functions structure, so that the DC programming approach of Section III-B cannot be used here. To deal with this issue, this section proposes instead the concave approximation of $b_k^n(\mathbf{P}_k)$ by $\tilde{b}_k^n(\mathbf{P}_k|\overline{\mathbf{P}}_k)$ around a given point $\overline{\mathbf{P}}_k$ as shown in (35) on the bottom of the next page. This approximation is based on the first-order Taylor expansion of the first log-term in (33) in the variables $\{\mathbf{p}_k^j\}$, together with the first-order Taylor expansion of the second log-term in the (convex) quadratic terms $\{\mathbf{p}_k^{j,\mathcal{H}} \mathbf{A}_k^n \mathbf{p}_k^j\}$. Notice that for the differentiation of a real function with respect to a complex variable, the rules of Wirtinger are used (see e.g. [37]). In addition, note that a first-order approximation of the second log-term in the variables $\{\mathbf{p}_k^j\}$ is not done here. It would in fact result in too much linearization and a very inaccurate approximation of (33).

Since approximation (35) is neither a lower bound nor an upper bound of $b_k^n(\mathbf{P}_k)$, a trust region approach is adopted to control the approximation accuracy. More specifically, an iterative method (listed in Alg. 3) is proposed for solving (31) that generates a sequence of iterates $\{\mathbf{P}_k(l)\}$, by relying on (35) that models the non-convex term in (32b) in a sufficiently small neighborhood of $\mathbf{P}_k(l)$, also known as the trust region [38]. Given this trust region-based model, the precoder matrices are updated by solving the following small-scale convex problem:

$$\mathbf{P}_k(l+1) = \underset{\{t_k^g\}, \mathbf{P}_k}{\arg\max} \sum_g \alpha_g t_k^g \tag{36a}$$

$$\text{s.t. } \tilde{b}_k^n\left(\mathbf{P}_k | \mathbf{P}_k(l)\right) \geq t_k^g, \quad \forall g, n \in \mathcal{G}_g \tag{36b}$$

$$\left\| \mathbf{p}_k^g - \mathbf{p}_k^g(l) \right\|_2^2 \leq \Delta_k^l, \quad \forall g \tag{36c}$$

$$\left[\mathbf{P}_k \mathbf{P}_k^{\mathcal{H}}\right]_{nn} \leq P_k^{\text{mask}}, \quad \forall n \tag{36d}$$

where (36c) is the trust region constraint with $\Delta_k^l \geq 0$ denoting the trust region radius at iteration l. Instrumental here is the choice of the trust region radius, which should result in a sufficient increase of the objective value in (31a) compared to the objective value predicted by the model in (36a). That is, Δ_k^l should be chosen for all tones k such that

$$\frac{\sum_g \alpha_g \left(c_k^g(l+1) - c_k^g(l)\right)}{\sum_g \alpha_g \left(t_k^g(l+1) - t_k^g(l)\right)} \geq \rho, \tag{37}$$

with $0 < \rho < 1$ a pre-defined constant [38]. The numerator and denominator in (37) are called the actual and predicted increase, respectively. Note that the predicted increase is always non-negative, such that this trust region condition yields a monotonically non-decreasing objective value. See [38], [39] for elaborate trust region radius update schemes. In Alg. 3 the radius is iteratively decreased with a factor $\beta_1 \leq 1$ until (37) is satisfied. Moreover, the following convergence result for Alg. 3 is provided in the theorem below.

Algorithm 3 Rank-1 Precoding-based Multicasting

Initialize $\boldsymbol{\lambda} \succeq \mathbf{0}$, $\{\rho, \beta_1\} \in (0, 1)$, $\beta_2 > 1$, Δ_k, $\Delta_k^{\max}$, $\forall k$
Initialize $\mathbf{P}_k^{(l)}$, $\forall k$ with feasible values
Set $l = 0$
repeat
  Set $l = l + 1$
  for $k = 1 \cdots K$ **do**
    repeat
      Update $\mathbf{P}_k(l)$ by solving (36) at tone k
      Update $c_k^g(l) = \min_{n \in \mathcal{G}_G} \left\{ b_k^n(\mathbf{P}_k(l)) \right\}$, $\forall g$
      Update $\Delta_k^l = \beta_1 \Delta_k^l$
    until condition (37) is satisfied
    Set $\Delta_k^{l+1} = \min(\beta_2 \Delta_k^l, \Delta_k^{\max})$
  end
until $\left| \sum_k \sum_g \alpha_g \left(c_k^g(l) - c_k^g(l-1) \right) \right| < \epsilon$

Theorem 2: The sequence $\{\mathbf{P}_k^{(l)}\}$ generated by Alg. 3 has a monotonically non-decreasing objective value and converges to a stationary point of problem (31).

Proof: The monotonically non-decreasing objective value convergence is easily established by the trust region condition in (37) that is satisfied every iteration. To show convergence to a stationary point, a similar proof outline as in [38, Chapter 12] is followed. For completeness, a full proof is provided in Appendix C. □

The computational complexity of Alg. 3 is dominated by the computation of the precoder matrices of size $N \times G$ for every iteration in optimization problem (36) for all tones. For each tone, problem (36) corresponds to quadratically constrained linear program (QCLP), which can be equivalently reformulated as a SOCP, and thus may be efficiently solved using the interior-point method in standard solvers such as CVX [30]. In terms of worst-case complexity, this requires $\mathcal{O}(\sqrt{NG}\log(1/\epsilon))$ iterations for an ϵ-accurate solution, with an approximate per-iteration complexity of $\mathcal{O}((NG)^3)$ [28]. Summing up, the total complexity of Alg. 3 is $\mathcal{O}(I_1\, K (NG)^{3.5} \log(1/\epsilon))$, with I_1 the number of SCA outer iterations (including the unsuccessful trials when (37) is not satisfied). It is remarked that the number of unsuccessful trials can be limited by efficiently keeping track of the trust region radii. Typically, 10-20 iterations are sufficient for Alg. 3 to converge as shown in Section V.

Moreover, note that Alg. 3 has a significantly lower per-iteration computational complexity than Alg. 2. This results mainly from the use of G precoder vectors of length N (i.e. rank-one precoding), instead of using G covariance matrices of size $N \times N$ (i.e. full-rank precoding). This yields a per-iteration computational complexity of $\mathcal{O}((NG)^3)$ for Alg. 3 versus $\mathcal{O}((N^2G)^3)$ for Alg. 2, since the computational complexity of interior-point methods scales (worst-case) cubically with the number of optimization variables.

B. RANK-ONE ZF-PRECODING

Inspired by [22] to further reduce the computational complexity, this section considers rank-one ZF-precoding for multi-group multicasting, based on the block-diagonalizing ZF precoder for downstream unicasting [23], [24]. In order to cancel all inter-group interference, the premise is that the precoder vector $\mathbf{p}_k^g$ of group g on tone k should lie in the null-space of

$$\tilde{\mathbf{H}}_k^{-g} \triangleq \left[\mathbf{H}_k^1, \ldots, \mathbf{H}_k^{g-1}, \mathbf{H}_k^{g+1}, \ldots, \mathbf{H}_k^G \right]^{\mathcal{H}} \tag{38}$$

with $\mathbf{H}_k^g$ a channel matrix of size $N \times |\mathcal{G}_g|$ containing the $|\mathcal{G}_g|$ channel vectors associated with group g on tone k. An efficient way to obtain these unitary null-space bases, is to compute the QR decomposition of the Hermitian conjugated $\tilde{\mathbf{H}}_k^{-g}$

$$(\tilde{\mathbf{H}}_k^{-g})^{\mathcal{H}} \overset{\text{QRD}}{=} \left[\tilde{\mathbf{Q}}_k^g, \mathbf{Q}_k^g \right] \begin{bmatrix} \mathbf{R}_k^g \\ \mathbf{0} \end{bmatrix} \tag{39}$$

where $\mathbf{Q}_k^g$ is an $N \times |\mathcal{G}_g|$ unitary null-space basis of $\tilde{\mathbf{H}}_k^{-g}$, i.e., $\mathbf{H}_k^{j,\mathcal{H}} \mathbf{Q}_g = \mathbf{0}$ for $j \neq g$. As a result, the ZF precoding matrix $\mathbf{P}_k$ may now be expressed as

$$\mathbf{P}_k = \left[\mathbf{Q}_k^1 \mathbf{v}_k^1, \ldots, \mathbf{Q}_k^G \mathbf{v}_k^G \right], \quad \forall k \tag{40}$$

where $\mathbf{v}_k^g$ is a complex vector of length $|\mathcal{G}_g|$.

Hence, defining $\hat{\mathbf{h}}_k^n \triangleq \mathbf{h}_k^{n,\mathcal{H}} \mathbf{Q}_k^{g_n}$, the design problem comes down to WSGR maximization in the inner precoder vectors $\{\mathbf{v}_k^g\}$ under per-line power constraints:

$$\underset{\{\mathbf{v}_k^g\}}{\text{maximize}} \sum_g \alpha_g \min_{n \in \mathcal{G}_g} \left\{ \log\left(1 + |\hat{\mathbf{h}}_k^n \mathbf{v}_k^g|^2\right) \right\} \tag{41a}$$

$$\text{s.t.} \left[\mathbf{P}_k \mathbf{P}_k^{\mathcal{H}} \right]_{nn} \leq P_k^{\text{mask}}, \quad \forall k, n. \tag{41b}$$

Notwithstanding there is no inter-group interference, problem (41) is still non-convex due to the quadratic term inside the log-function. Problem (41) can be tackled with a similar iterative trust-region method as in Section IV-A, by using in each iteration a first-order Taylor expansion of the log-term around the previous operating point $\{\overline{\mathbf{v}}_k^g\}$, i.e.,

$$\log\left(1 + |\hat{\mathbf{h}}_k^n \mathbf{v}_k^{g_n}|^2\right) \approx \log\left(1 + |\hat{\mathbf{h}}_k^n \overline{\mathbf{v}}_k^g|^2\right) + \frac{2}{1 + |\hat{\mathbf{h}}_k^n \overline{\mathbf{v}}_k^{g_n}|^2} \left(\text{Re}\left\{ \overline{\mathbf{v}}_k^{g_n,\mathcal{H}} \hat{\mathbf{h}}_k^n \hat{\mathbf{h}}_k^{n,\mathcal{H}} \left(\mathbf{v}_k^{g_n} - \overline{\mathbf{v}}_k^{g_n} \right) \right\} \right), \quad \forall n. \tag{42}$$

$$\tilde{b}_k^n(\mathbf{P}_k | \overline{\mathbf{P}}_k) = \log\left(\overline{X}_k^n + \overline{S}_k^n\right) + \frac{2}{\overline{X}_k^n + \overline{S}_k^n} \left(\sum_j \text{Re}\left\{ \overline{\mathbf{p}}_k^{j,\mathcal{H}} \mathbf{A}_k^n \left(\mathbf{p}_k^j - \overline{\mathbf{p}}_k^j \right) \right\} \right) - \log\left(\overline{X}_k^n\right) - \frac{1}{\overline{X}_k^n} \sum_{j \neq g_n} \left(\mathbf{p}_k^{j,\mathcal{H}} \mathbf{A}_k^n \mathbf{p}_k^j - \overline{\mathbf{p}}_k^{j,\mathcal{H}} \mathbf{A}_k^n \overline{\mathbf{p}}_k^j \right) \tag{35}$$

Further details are omitted for brevity. Assuming the same number of users in each group, i.e. $|\mathcal{G}| = |\mathcal{G}_g|, \forall g$, with $N \geq |\mathcal{G}|$, the total complexity reduces to $\mathcal{O}(I_1\ K(|\mathcal{G}|G)^{3.5}\log(1/\epsilon))$, with I_1 the number of SCA outer iterations. Due to the ZF constraints, the number of SCA iterations I_1 is typically smaller than with general rank-one precoding. Moreover, heuristic methods to compute the inner precoder vectors $\{\mathbf{v}_k\}$ are possible. For instance, the heuristic method in [22] for the (single-tone) max-min-fairness problem with a sum-power constraint can be used for initialization (with simply scaling of the precoder vectors to satisfy the per-line power constraints).

Remark 2: This rank-one ZF-precoding-based multicasting scheme is a generalization of the BD-based ZF precoder for downstream unicasting in [40]. In case of N single-user groups, (41) turns into a smooth yet still non-convex problem. However, in this case it may be solved by convex SDR, which is tight since [40] provides a Lagrange dual decomposition approach with an optimal rank-one solution. Such a SDR approach is not tight in general for multicasting.

V. G.FAST CABLE BINDER SIMULATION

In this section, a cable binder is simulated consisting of 10 lines with a length of 80 m, for the downstream G.fast 212 MHz profile and using various multicasting schemes. The channel matrices have been obtained by measurements of a single cable binder. Following the G.fast recommendation [2], the per-line ATP constraints are 8 dBm while the per-tone PSD spectral masks are obtained from [41] ranging from −65 dBm/Hz to −79 dBm/Hz. The tone spacing Δ_f is 51.75 kHz and the noise PSD is assumed to be −140 dBm/Hz. The symbol rate is 48 kHz. In these simulations, the obtained bit loadings are capped at 14 bits and the ATP constraints are always observed to be inactive. The following schemes below are considered.

- *ZF-Precoding-Based Unicasting (ZF-UC)*: is the baseline unicasting scheme. It uses the ZF precoder matrix, whereas the power allocation is optimized for WSGR maximization, using a similar Langrange dual decomposition method as in Section III.
- *Rank-One ZF-Precoding-Based Multicasting* (Rank-1 ZF-MC): as presented in Section IV-B. The initialization is based on a scaling of the inner precoder vectors obtained by the heuristic successive precoding method in [22, Alg. 3] to satisfy the per-line power constraints.
- *Rank-One Precoding-Based Multicasting* (Rank-1 MC): as presented in Section IV-A. The same initialization procedure as in Rank-1 ZF-MC is used.
- *Full-Rank MC*: as presented in Section III-B. To speed up convergence, it is initialized with the solution of Rank-1 MC.

A. SINGLE-GROUP MULTICASTING

In this subsection, the 10-line G.fast cable binder is considered to be a single-group. The numerical results in Table 1 show that all multicasting schemes achieve a significant gain over the baseline unicasting scheme. The multicasting capacity is provided by the Full-Rank MC with channel coding across tones (Section III-A). However, to the best of our knowledge, this capacity cannot be practically achieved by any physical-layer multicasting scheme. Moreover, the corresponding optimization method (Alg. 1) is very computationally complex due to the full-rank transmit covariance matrices as optimization variables in combination with the iterative Lagrange multiplier search. By contrast, rank-one precoding-based multicasting provides practically achievable date-rates and significantly lowers the computational complexity (note that ZF-Rank-1 and Rank-1 MC are the same for single-group multicasting). Moreover, the results show that the performance gap between rank-one precoding and the multicasting capacity is rather limited. In addition, the SDR of problem (31) for the single-group case is included, leading to a convex SDP which yields a global upper bound of (31) on every tone separately [10]. This is equivalent to problem (19) with the summation over the tones shifted outside the min-function, i.e., full-rank precoding-based multicasting with independent per-tone channel coding (Full-Rank MC-PTCC). The small gap between Full-Rank MC-PTCC and Rank-1 MC indicates that Alg. 3 efficiently solves the non-convex problem (31) for the single-group case. Finally, standard unicasting (i.e. ZF-UC) is near-optimal for low-frequency channels, which are diagonally dominant (see Fig. 2). By consequence, the relative gain of multicasting

TABLE 1. Single-group multicasting scenario.

Scheme	Start. Freq. 2.2 MHz Single Group-Rate [Mbps]	[%]	Start. Freq. 40 MHz Single Group-Rate [Mbps]	[%]
ZF-UC	1414	70.6	920	61.1
Rank-1 MC	1967	98.2	1473	97.7
Full-Rank MC-PTCC	1986	99.2	1492	98.9
Full-Rank MC	2002	100	1508	100

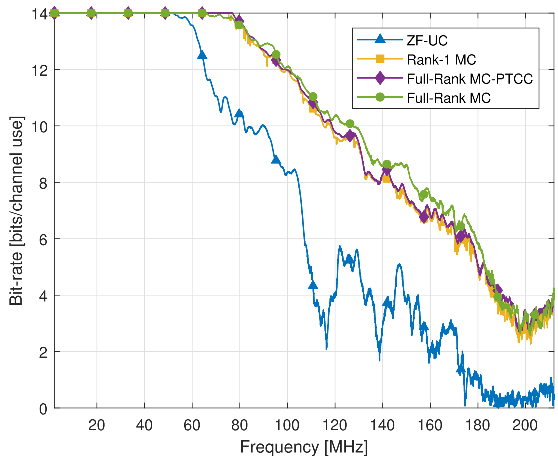


FIGURE 2. The minimum bit-rate across the frequency for the single-group multicasting G.fast scenario. For ZF-UC and Full-Rank MC, this corresponds to the mean user bit-rate.

TABLE 2. Multi-group multicasting – Scenario A.

Scheme	Start. Freq. 2.2 MHz Avg. Group-Rate [Mbps]	[%]	Start. Freq. 40 MHz Avg. Group-Rate [Mbps]	[%]
ZF-UC	1448	81.9	954	76.6
ZF-Rank-1 MC	1661	94.0	1167	91.9
Rank-1 MC	1730	97.9	1236	97.1
Full-Rank MC	1768	100	1274	100

TABLE 3. Multi-group multicasting – Scenario B.

Scheme	Start. Freq. 2.2 MHz Avg. Group-Rate [Mbps]	[%]	Start. Freq. 40 MHz Avg. Group-Rate [Mbps]	[%]
ZF-UC	1521	85.3	1027	79.7
ZF-Rank-1 MC	1627	91.3	1134	86.7
Rank-1 MC	1740	97.6	1247	95.6
Full-Rank MC	1782	100	1289	100

is typically larger in case of a higher starting frequency (e.g. to enforce spectral compatibility with VDSL2 35MHz in the same cable binder)

B. MULTI-GROUP MULTICASTING

A first considered multi-group scenario (referred to as scenario A) is obtained by dividing the 10-line G.fast cable binder into $G = 4$ multicast groups with two or three users per group. A second multi-group scenario (referred to as scenario B) consists of a single six-user multicast group, together with four unicast users (i.e. four single-user groups). In contrast to scenario A, scenario B exhibits strong non-convexity with possibly very sub-optimal solutions, due to the unbalanced group sizes. For instance, in case all five groups have the same unity weight $\{\alpha_g = 1\}$, the multicasting schemes tend to shut down the large multicasting group in favor of the single-user groups. To counter this, the weight of the large multicasting group is increased to $\alpha_1 = 3$, which is fair as this group represents more users than the single-user groups. The numerical results of scenario A are summarized in Table 2 for the case with unit group weights ($\alpha_g = 1, \forall g$). The numerical results of scenario B are summarized in Table 3.

The numerical results show that also for the multi-group case the G.fast channels admit a significant multicasting gain, although naturally smaller than in the single-group case. Nonetheless, the average bit-rate achieved by each group with Rank-1 MC is increased by more than 200 Mbps in both scenarios over the baseline unicasting scheme. Further, the performance gap between Rank-1 MC and Full-rank MC (the latter providing an information theoretic upper bound) is rather small. Besides being practically achievable, Rank-1 MC also has a much lower computational complexity than Full-rank MC, since its corresponding problem (31) is fully decoupled across the tones and only requires the computation of rank-one precoder matrices (instead of full-rank covariance matrices, see Section IV-A). The gain of Rank-1 MC over Rank-1 ZF-MC is primarily achieved in the higher frequency part of the spectrum (see Fig. 3 and 4), where allowing some residual inter-group interference in exchange for useful signal power gain yields a net SINR gain.

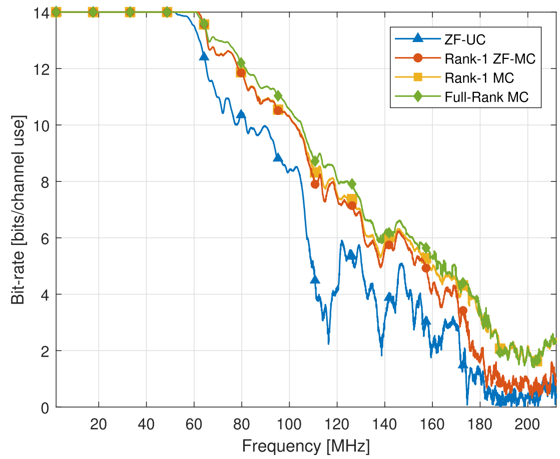


FIGURE 3. The average group bit-rate across the frequency for the multi-group multicasting G.fast scenario A.

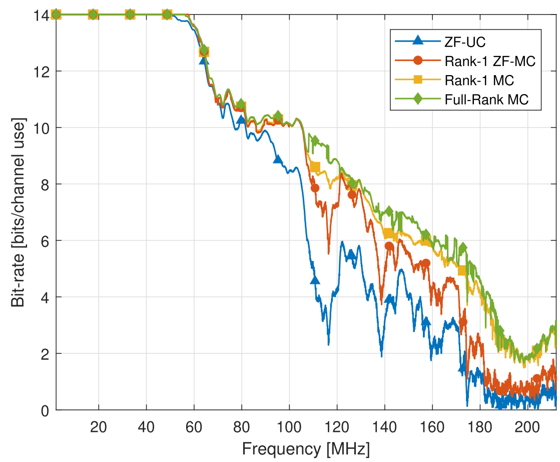


FIGURE 4. The average group bit-rate across the frequency for the multi-group multicasting G.fast scenario B.

The convergence of the proposed iterative trust-region methods (i.e. Rank-1 and ZF-Rank-1 MC) is shown in Fig. 5 for scenario A and B. In each iteration, if condition (37) is not satisfied, the objective from the previous successful iteration is used. Fig. 5 shows a fast convergence, with a relative objective increase smaller than 10^{-3} after ten iterations in both scenarios. For comparison, the SCA-based algorithm proposed in [21] for minimum group rate maximization is adopted here for WSGR maximization and included in Fig. 5. This method involves a non-linear constraint in each iteration, which for the WSGR case however may be converted equivalently into a geometric mean objective (which in turn

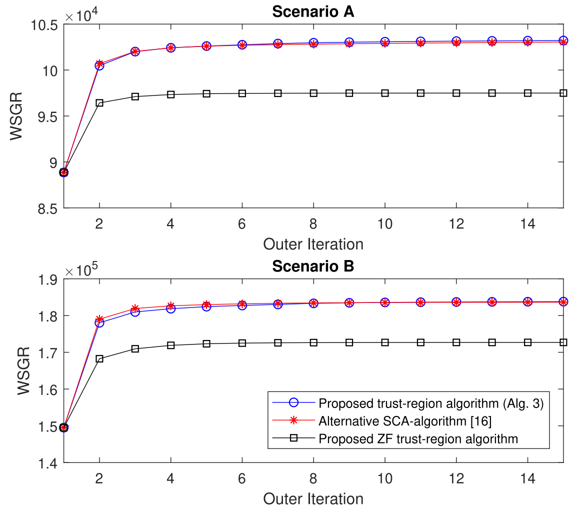


FIGURE 5. Convergence of the proposed iterative trust-region methods for the multi-group multicasting G.fast scenario A and B.

may be transformed into a system of SOC constraints [42]). Both SCA methods exhibit a similar convergence speed and achieve almost the same WSGR (only on some high frequency tones a different local optimum is obtained), which demonstrates the viability and stability of SCA for rank-one precoding-based multicasting in G.fast DSL networks.

VI. CONCLUSION

This paper studied physical-layer multicasting design for downstream G.fast DSL transmission, which corresponds to a multi-user multi-tone (i.e., multi-carrier) scenario. As an information-theoretic upper bound, the transmit covariance matrices are assumed to be full-rank in combination with joint channel coding across the tones. For a single group, this corresponds to the multicasting capacity. Maximizing the minimum total bit-rate leads to a non-linear convex SDP, which is solved with Lagrange dual decomposition, in order to decouple the problem across tones. This approach is extended for maximizing the WSGR in a multi-group multicasting scenario, by relying on DC programming. In addition, rank-one single-stream precoding with independent per-tone channel coding is considered as a practical multicasting scheme. To compute the precoding matrices in this scheme, instead of relying on computationally complex SDR, a SCA-based trust-region algorithm is developed. Finally, simulations of a G.fast cable binder are provided to show that the practical multicasting schemes operate close to the information-theoretic multicasting upper bound.

APPENDIX A
EUCLIDEAN PROJECTION OF θ ONTO THE CONSTRAINT SET

In each iteration i of the dual decomposition algorithm, the subgradient update vector $\boldsymbol{\theta}^{(i+1)} = [\bar{\theta}_1, \ldots, \bar{\theta}_N]$ needs to be jointly Euclidean projected onto two constraint sets: $C^1 = \{\boldsymbol{\theta} | \mathbf{1}^T\boldsymbol{\theta} = 1\}$ and $C^2 = \{\boldsymbol{\theta} | \boldsymbol{\theta} \succeq \mathbf{0}\}$. This corresponds to the following problem:

$$\begin{aligned} &\underset{\{\theta_n\}}{\text{minimize}} \ \frac{1}{2}\sum_n \left(\theta_n - \bar{\theta}_n\right)^2 \\ &\text{s.t.} \ \sum_n \theta_n = 1, \theta_n \geq 0, \quad \forall n, \end{aligned} \tag{43}$$

The (partial) Lagrangian of the above problem is

$$\mathcal{L}(\boldsymbol{\theta}, \mu) = \frac{1}{2}\sum_n \left(\theta_n - \bar{\theta}_n\right)^2 + \mu\left(\sum_n \theta_n - 1\right). \tag{44}$$

It is observed that for fixed μ, minimizing the $\mathcal{L}(\boldsymbol{\theta}, \mu)$ subject to non-negative $\{\theta_n\}$ admits the following closed-form solution:

$$\theta_n = \left[\bar{\theta}_n - \mu\right]^+. \tag{45}$$

Hence, the optimal μ can be derived by enforcing

$$\sum_n \left[\bar{\theta}_n - \mu\right]^+ = 1, \tag{46}$$

which requires $\mathcal{O}(N\log(N))$ operations. More specifically, first sorting the $\{\bar{\theta}_n\}$ values, which takes $\mathcal{O}(N\log(N))$ operations. Suppose that $\bar{\theta}_1 \leq \bar{\theta}_2 \leq \ldots, \leq \bar{\theta}_N$. Then testing whether the desired μ belongs to the interval $[\bar{\theta}_n, \bar{\theta}_{n+1}]$ reduces the problem to a univariate linear equation.

APPENDIX B
PROOF OF THEOREM 1

Alg. 2 is a special case of the non-smooth successive upper bound minimization (SUM) algorithm presented in [34], with results that are also valid for the case with complex variables. Let $f(\mathbf{C}^{(l)})$ and $m(\mathbf{C}|\mathbf{C}^{(l-1)})$ denote the objective value in (19a) and (23a), respectively, with $m(\mathbf{C}|\mathbf{C}^{(l-1)})$ referred to as the model. Observe then that $f(\mathbf{C})$ and $m(\mathbf{C}|\mathbf{C}^{(l-1)})$ are composite functions, with $\sum_g \alpha_g \min\{\cdot\}$ a concave function and the $\{b_k^n(\mathbf{C}_k)\}$ and $\{\tilde{b}_k^n(\mathbf{C}_k|\mathbf{C}_k^{(l-1)})\}$ continuously differentiable functions. Additionally, the power constraints (19b) and (19c) form a convex and compact feasible region $\mathcal{C}$ for the transmit covariance matrices $\mathbf{C}$. Then the following conditions hold:

1) The model $m(\mathbf{C}|\overline{\mathbf{C}})$ is continuous in $(\mathbf{C}, \overline{\mathbf{C}}) \in \mathcal{C} \times \mathcal{C}$.
2) The model is a tight and global lower bound for the objective function, i.e., for all $\overline{\mathbf{C}} \in \mathcal{C}$:
$$\begin{aligned} &f(\overline{\mathbf{C}}) = m(\overline{\mathbf{C}}|\overline{\mathbf{C}}) \\ &f(\mathbf{C}) \geq m(\mathbf{C}|\overline{\mathbf{C}}) \text{ for any set } \mathbf{C} = \{\mathbf{C}_k \succeq \mathbf{0}\}. \end{aligned}$$
3) The directional derivative of the model and the objective function are equal in every point $\mathbf{C} = \overline{\mathbf{C}} \in \mathcal{C}$.

Hence, the monotonically non-decreasing convergence of $f(\mathbf{C}^{(l)})$ is established by observing that

$$f(\mathbf{C}^{(l)}) \geq m(\mathbf{C}^{(l)}|\mathbf{C}^{(l-1)}) \geq m(\mathbf{C}^{(l-1)}|\mathbf{C}^{(l-1)}) = f(\mathbf{C}^{(l-1)}) \quad \text{for } l = 1, 2, \ldots \tag{47}$$

due to condition 2 and the optimality of solving (23). Moreover, since the generated sequence $\{\mathbf{C}^{(l)}\}$ is always in the bounded set $\mathcal{C}$, every subsequence (of which there exists at least one) converges to a certain limit point $\mathbf{C}^*$. In combination with (47) this implies that

$$f(\mathbf{C}^*) = m(\mathbf{C}^*|\mathbf{C}^*) \geq m(\mathbf{C}|\mathbf{C}^*) \text{ for any } \mathbf{C} \in \mathcal{C}, \quad (48)$$

and hence that $\mathbf{C}^*$ is a stationary point of (24), and thus equivalently also of (19) due to condition 3 [34].

APPENDIX C
PROOF OF THEOREM 2

Define the objective function of problem (31) at tone k as $f_k(\mathbf{P}_k) \triangleq \sum_g \alpha_g \min_{n\in\mathcal{G}_g}\{b_k^n(\mathbf{P}_k)\}$. As each tone k is handled independently, the tone index in this appendix is dropped for conciseness. The monotonically non-decreasing convergence of $f(\mathbf{P}^{(l)})$ is established by the trust region condition in (37) that is satisfied in each iteration. To show convergence to a stationary point, a similar proof outline as in [38, Chapter 12] is followed. Observe that $f(\mathbf{P})$ is a composite function where $\sum_g \alpha_g \min\{\cdot\}$ is a concave function and the $\{b_k^n(\mathbf{P}_k)\}$ are continuously differentiable; and that the power constraints (31c) form a convex and compact feasible region $\mathcal{P}$ for the precoding matrix $\mathbf{P}$, such that $f(\mathbf{P})$ is bounded from above. Then the following conditions hold [38], [43]:

1) The objective function $f(\mathbf{P})$ is locally Lipschitz continuous and regular in $\mathbf{P} \in \mathcal{P}$.
2) The model $m(\mathbf{P}|\overline{\mathbf{P}})$ is locally Lipschitz continuous and regular in $\mathbf{P} \in \mathcal{P}$ for all $\overline{\mathbf{P}} \in \mathcal{P}$, and continuous in $(\mathbf{P}, \overline{\mathbf{P}})$.
3) The objective function and model have the same value when $\mathbf{P} = \overline{\mathbf{P}}$, i.e., $f(\overline{\mathbf{P}}) = m(\overline{\mathbf{P}}|\overline{\mathbf{P}})$ for all $\overline{\mathbf{P}} \in \mathcal{P}$.
4) The directional derivative of the model and the objective function are equal in every point $\mathbf{P} = \overline{\mathbf{P}} \in \mathcal{P}$.

Further, based on [43], consider the following specific first-order criticality measure for problem (31) in $\overline{\mathbf{P}} \in \mathcal{P}$, without loss of generality:

$$\pi_\mu\left(\overline{\mathbf{P}}\right) \triangleq \underset{\mathbf{P}\in\mathcal{P}}{\text{maximize}}\ m(\mathbf{P}|\overline{\mathbf{P}}) - f(\overline{\mathbf{P}})$$
$$\text{s.t. } \|\mathbf{p}_k^g - \overline{\mathbf{p}}_k^g\|_2^2 \leq \mu, \quad \forall g. \quad (49)$$

Note that $\pi_\mu\left(\overline{\mathbf{P}}\right)$ is always positive for any $\mu > 0$ and non-decreasing in μ for all $\mu > 0$. Moreover, $\pi_\mu\left(\overline{\mathbf{P}}\right) = 0$ if and only if $\overline{\mathbf{P}}$ is a stationary point of (31). This means that there exists no feasible direction in which the model $m(\mathbf{P}|\overline{\mathbf{P}})$ (which is a first-order approximation of $f(\mathbf{P})$) increases. The condition that Alg. 3 converges to the set of stationary points hence corresponds to

$$\lim_{l\to\infty} \pi_1\left(\mathbf{P}^{(l)}\right) = 0. \quad (50)$$

Condition (50) can be proven to hold by contradiction. As the generated sequence $\{\mathbf{P}^{(l)}\}$ by Alg. 3 is in the bounded set $\mathcal{P}$, every subsequence must converge to a certain limit point $\mathbf{P}^*$. Suppose now that such a limit point is not a stationary point, i.e. $\pi_1\left(\mathbf{P}^*\right) > 0$. Then, it can be shown that there exist strictly positive trust region radii $\Delta_l \geq \Delta_{\min} > 0$ which satisfy (37) for all iterations l of that subsequence, by using conditions 1 to 4 [38, Th. 11.2.3/11.2.4]. This gives for every iteration l

$$f(\mathbf{P}^{(l+1)}) - f(\mathbf{P}^{(l)}) \geq \rho\left(m(\mathbf{P}^{(l+1)}|\mathbf{P}^{(l)}) - f(\mathbf{P}^{(l)})\right) \quad (51)$$
$$= \rho\pi_{\Delta_l}\left(\mathbf{P}^{(l)}\right), \quad (52)$$

where in case $\Delta_l \geq 1$, we have that

$$f(\mathbf{P}^{(l+1)}) - f(\mathbf{P}^{(l)}) \geq \rho\pi_{\Delta_l}\left(\mathbf{P}^{(l)}\right) \geq \rho\pi_1\left(\mathbf{P}^{(l)}\right) > 0. \quad (53)$$

Otherwise, in case $\Delta_l < 1$, we have that

$$f(\mathbf{P}^{(l+1)}) - f(\mathbf{P}^{(l)}) \geq \rho\pi_{\Delta_l}\left(\mathbf{P}^{(l)}\right) \geq \rho\Delta_l\pi_1\left(\mathbf{P}^{(l)}\right)$$
$$\geq \rho\Delta_{\min}\pi_1\left(\mathbf{P}^{(l)}\right) > 0, \quad (54)$$

where (54) is valid due to $\pi_x(\cdot)/x$ being non-increasing in $x > 0$ [43, Lemma 2.1(iv)]. This means that the objective value increases in every iteration as long as a stationary point is not approached, i.e., $f(\mathbf{P}^{(l+1)}) > f(\mathbf{P}^{(l)}), \forall l$, such that the subsequence $f(\mathbf{P}^{(l)})$ grows to infinity when l goes to infinity. This contradicts with $f(\mathbf{P})$ being bounded from above due to the power constraints. Hence, $\mathbf{P}^*$ must be a stationary point with $\pi_1\left(\mathbf{P}^*\right) = 0$.

ACKNOWLEDGMENT

This research work is presented in the Ph.D. dissertation of [1]. The scientific responsibility is assumed by its authors.

REFERENCES

[1] W. Lanneer, "Vectoring design optimization in ultra-broadband DSL networks," Ph.D. dissertation, KU Leuven, Leuven, Belgium, Feb. 2019.
[2] *Fast Access to Subscriber Terminals (FAST)—Physical Layer Specification*, document Rec. ITU-T G.9701-Amendment 3, Apr. 2017.
[3] R. Strobel, R. Stolle, and W. Utschick, "Wideband modeling of twisted-pair cables for MIMO applications," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2013, pp. 2828–2833.
[4] 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.
[5] R. Strobel, *Precoding Optimization for Copper Line Networks*. Cham, Switzerland: Springer, 2019, pp. 33–101. doi: 10.1007/978-3-319-91560-9_3.
[6] M. Hekrdla, A. Matera, U. Spagnolini, and W. Wang, "Per-line power controlled lattice-reduction aided zero-forcing precoding for G.fast downstream," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–6.
[7] T. Quek, M. Peng, O. Simeone, and W. Yu, *Cloud Radio Access Networks: Principles, Technologies, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2017.
[8] B. Dai, Y.-F. Liu, and W. Yu, "Optimized base-station cache allocation for cloud radio access network with multicast backhaul," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 8, pp. 1737–1750, Jun. 2018.
[9] J. Cioffi, P. Chow, and K. Kerpez. (Nov. 2017). DSL's on going efficient enablement of the Internet continues. Long live DSL! ASSIA. [Online]. Available: http://assiainc.wpengine.com/wp-content/uploads/2018/05/Long-Live-DSL-Nov-2017.pdf
[10] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.
[11] S. X. Wu, W.-K. Ma, and A. M.-C. So, "Physical-layer multicasting by stochastic transmit beamforming and Alamouti space-time coding," *IEEE Trans. Signal Process.*, vol. 61, no. 17, pp. 4230–4245, Sep. 2013.
[12] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang, "Semidefinite Relaxation of Quadratic Optimization Problems," *IEEE Signal Process. Mag.*, vol. 27, no. 3, pp. 20–34, May 2010.

[13] E. Karipidis, N. Sidiropoulos, and Z.-Q. Luo, "Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1268–1279, Mar. 2008.

[14] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.

[15] L.-N. Tran, M. F. Hanif, and M. Juntti, "A conic quadratic programming approach to physical layer multicasting for large-scale antenna arrays," *IEEE Signal Process. Lett.*, vol. 21, no. 1, pp. 114–117, Jan. 2014.

[16] C. Lu and Y.-F. Liu, "An efficient global algorithm for single-group multicast beamforming," *IEEE Trans. Signal Process.*, vol. 65, no. 14, pp. 3761–3774, Jul. 2017.

[17] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Multicast multigroup beamforming for per-antenna power constrained large-scale arrays," in *Proc. IEEE SPAWC*, Jun. 2015, pp. 271–275.

[18] O. Mehanna, K. Huang, B. Gopalakrishnan, A. Konar, and N. D. Sidiropoulos, "Feasible point pursuit and successive approximation of non-convex QCQPs," *IEEE Signal Process. Lett.*, vol. 22, no. 7, pp. 804–808, Jul. 2015.

[19] M. Kaliszan, E. Pollakis, and S. Stańczak, "Multigroup multicast with application-layer coding: Beamforming for maximum weighted sum rate," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2012, pp. 2270–2275.

[20] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Sum rate maximizing multigroup multicast beamforming under per-antenna power constraints," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2014, pp. 3354–3359.

[21] G. Venkatraman, A. Tölli, M. Juntti, and L.-N. Tran, "Multigroup multicast beamformer design for MISO-OFDM with antenna selection," *IEEE Trans. Signal Process.*, vol. 65, no. 22, pp. 5832–5847, Nov. 2017.

[22] M. Sadeghi, L. Sanguinetti, R. Couillet, and C. Yuen, "Reducing the computational complexity of multicasting in large-scale antenna systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2963–2975, May 2017.

[23] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[24] L.-U. Choi and R. D. Murch, "A transmit preprocessing technique for multiuser MIMO systems using a decomposition approach," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 20–24, Jan. 2004.

[25] N. Jindal and Z. Luo, "Capacity limits of multiple antenna multicast," in *Proc. Int. Symp. Inf. Theory*, Jul. 2006, pp. 1841–1845.

[26] D. Tse and P. Viswanath, *Fundamentals Wireless Communication*. New York, NY, USA: Cambridge Univ. Press, 2005.

[27] Y. Huang and S. Zhang, "Complex matrix decomposition and quadratic programming," *Math. Oper. Res.*, vol. 32, no. 3, pp. 758–768, 2007.

[28] Y. Ye, *Interior Point Algorithms: Theory and Analysis*. New York, NY, USA: Wiley, 1997.

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge Univ. Press, 2004.

[30] M. Grant and S. Boyd. (Mar. 2014). *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.1*. [Online]. Available: http://cvxr.com/cvx

[31] S. Boyd and J. Park, "Subgradient method," Stanford Univ., Stanford, CA, USA, Class Notes, May 2014. [Online]. Available: https://web.stanford.edu/class/ee364b/lectures/subgrad_method_notes.pdf

[32] V. L. Nir, M. Moonen, J. Verlinden, and M. Guenach, "Full vectoring optimal power allocation in xDSL channels under per-modem power constraints and spectral mask constraints," *IEEE Trans. Commun.*, vol. 57, no. 1, pp. 194–202, Jan. 2009.

[33] M. H. M. Costa, "Writing on dirty paper (corresp.)," *IEEE Trans. Inf. Theory*, vol. 29, no. 5, pp. 439–441, May 1983.

[34] 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, 2013. [Online]. Available: https://doi.org/10.1137/120891009

[35] J. Verdyck, W. Lanneer, P. Tsiaflakis, W. Coomans, P. Patrinos, and M. Moonen, "Optimal dynamic spectrum management algorithms for multi-user full-duplex DSL," *IEEE Access*, 2019.

[36] A. Suls, J. Verdyck, M. Moonen, and M. Moeneclaey, "A modified SNR gap approximation for resource allocation in LDPC-coded multicarrier systems," ESAT-STADIUS, Katholieke Univ. Leuven, Leuven, Belgium, Internal Rep. 19-27, 2019.

[37] R. Hunger, "An introduction to complex differentials and complex differentiability," Tech. Univ. München, München, Munich, Germany, Tech. Rep. TUM-LNS-TR-07-06, 2007.

[38] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust Region Methods*. Philadelphia, PA, USA: SIAM, 2000. [Online]. Available: https://epubs.siam.org/doi/abs/10.1137/1.9780898719857

[39] J. Nocedal and S. J. Wright, *Numerical Optimization*, 2nd ed. New York, NY, USA: Springer, 2006.

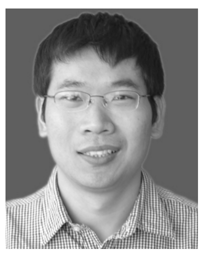
[40] R. Zhang, "Cooperative multi-cell block diagonalization with per-base-station power constraints," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1435–1445, Dec. 2010.

[41] *Fast Access to Subscriber Terminals (FAST)—Power Spectral Density Specification*, document Recommendation ITU-T G.9700, Apr. 2014.

[42] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, "Applications of second-order cone programming," *Linear Algebra Appl.*, vol. 284, nos. 1–3, pp. 193–228, Nov. 1998.

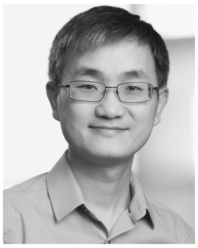
[43] Y. Yuan, "Conditions for convergence of trust region algorithms for non-smooth optimization," *Math. Programm.*, vol. 31, no. 2, pp. 220–228, Jun. 1985.

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 with Nokia Bell Labs, Antwerp, Belgium, where he is involved in MIMO communications, signal processing, and optimization for next-generation communication networks.



YA-FENG LIU (M'12–SM'18) received the B.Sc. degree in applied mathematics from Xidian University, Xi'an, China, in 2007, and the Ph.D. degree in computational mathematics from the Chinese Academy of Sciences (CAS), Beijing, China, in 2012. During his Ph.D. study, he was supported by the Academy of Mathematics and Systems Science (AMSS), CAS, to visit Professor Z.-Q. (Tom) Luo at the University of Minnesota, Twins Cities, from 2011 to 2012. After his graduation, he joined the Institute of Computational Mathematics and Scientific/Engineering Computing, AMSS, CAS, in 2012, where he became an Associate Professor, in 2018. His main research interests include nonlinear optimization and its applications to signal processing, wireless communications, and machine learning.

Dr. Liu was a recipient of the Best Paper Award from the IEEE International Conference on Communications (ICC), in 2011, the Best Student Paper Award from the International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), in 2015, the Chen Jingrun Star Award from AMSS, and the Science and Technology Award for Young Scholars from the Operations Research Society of China, in 2018. He has been an Editor of the IEEE Transactions on Wireless Communications, since 2019, and a Guest Editor of the *Journal of Global Optimization*, since 2016.



WEI YU (S'97–M'02–SM'08–F'14) received the B.A.Sc. degree in computer engineering and mathematics from the University of Waterloo, Waterloo, ON, Canada in 1997, and the M.S. and Ph.D. degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1998 and 2002, respectively. Since 2002, he has been with the Electrical and Computer Engineering Department, University of Toronto, Toronto, ON, Canada, where he is currently a Professor and holds the Canada Research Chair (Tier 1) in Information Theory and Wireless Communications. His main research interests include information theory, optimization, wireless communications, and broadband access networks.

Dr. Yu serves as the Second Vice President of the IEEE Information Theory Society, in 2019. He is a Fellow of the Canadian Academy of Engineering and a member of the College of New Scholars, Artists and Scientists, Royal Society of Canada. He was an IEEE Communications Society Distinguished Lecturer, from 2015 to 2016. He received the Steacie Memorial Fellowship, in 2015. He was a recipient of the IEEE Marconi Prize Paper Award in wireless communications, in 2019, the IEEE Communications Society Award for Advances in Communication, in 2019, the IEEE Signal Processing Society Best Paper Award, in 2017 and 2008, the Journal of Communications and Networks Best Paper Award, in 2017, the IEEE Communications Society Best Tutorial Paper Award, in 2015, the IEEE International Conference on Communications Best Paper Award, in 2013, the McCharles Prize for Early Career Research Distinction, in 2008, the Early Career Teaching Award from the Faculty of Applied Science and Engineering, University of Toronto, in 2007, and an Early Researcher Award from Ontario, in 2006. He served as the Chair for the Signal Processing for Communications and Networking Technical Committee of the IEEE Signal Processing Society, from 2017 to 2018. He is recognized as a Highly Cited Researcher. He is currently an Area Editor for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, from 2017 to 2020, and in the past, he served as an Associate Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY, from 2010 to 2013, and as an Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, from 2009 to 2011, and the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, from 2004 to 2007.

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.

He is a Fellow of EURASIP in 2018. He received the 1994 KU Leuven Research Council Award, the 1997 Alcatel Bell (Belgium) Award (with P. Vandaele), the 2004 Alcatel Bell (Belgium) Award (with R. Cendrillon), and was a 1997 Laureate of the Belgium Royal Academy of Science. He also received 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 Technical Committee on Signal Processing for Communications of the IEEE Signal Processing Society, and a President of the European Association for Signal Processing (EURASIP), from 2007 to 2008 and from 2011 to 2012. He has served as the Editor-in-Chief for the *EURASIP Journal on Applied Signal Processing*, from 2003 to 2005, 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*, *Integration*-the VLSI Journal, the *EURASIP Journal on Wireless Communications and Networking*, and the *EURASIP Journal on Advances in Signal Processing*.