

Multicast Multigroup Precoding and User Scheduling for Frame-Based Satellite Communications

Dimitrios Christopoulos, *Member, IEEE*, Symeon Chatzinotas, *Senior Member, IEEE*, and Björn Ottersten, *Fellow, IEEE*

Abstract—The present work focuses on the forward link of a broadband multibeam satellite system that aggressively reuses the user link frequency resources. Two fundamental practical challenges, namely the need to frame multiple users per transmission and the per-antenna transmit power limitations, are addressed. To this end, the so-called frame-based precoding problem is optimally solved using the principles of physical layer multicasting to multiple co-channel groups under per-antenna constraints. In this context, a novel optimization problem that aims at maximizing the system sum rate under individual power constraints is proposed. Added to that, the formulation is further extended to include availability constraints. As a result, the high gains of the sum rate optimal design are traded off to satisfy the stringent availability requirements of satellite systems. Moreover, the throughput maximization with a granular spectral efficiency versus SINR function, is formulated and solved. Finally, a multicast-aware user scheduling policy, based on the channel state information, is developed. Thus, substantial multiuser diversity gains are gleaned. Numerical results over a realistic simulation environment exhibit as much as 30% gains over conventional systems, even for 7 users per frame, without modifying the framing structure of legacy communication standards.

Index Terms—Broadband multibeam satellite systems, optimal linear precoding, sum rate maximization, multicast multigroup beamforming, per-antenna constraints.

I. INTRODUCTION AND RELATED WORK

AGGRESSIVE frequency reuse schemes have shown to be the most promising way towards spectrally efficient, high-throughput wireless communications. In this context, linear precoding, a transmit signal processing technique that exploits the offered spatial degrees of freedom of a multi-antenna transmitter, is brought into play to manage interferences. Such interference mitigation techniques and subsequently full frequency reuse configurations, are enabled by the availability of channel state information (CSI) at the transmitter.

In fixed broadband multibeam satellite communications (satcoms), the relatively slow channel variations facilitate the channel acquisition process. Therefore, such scenarios emerge

Manuscript received July 2, 2014; revised January 14, 2015 and April 3, 2015; accepted April 5, 2015. Date of publication April 21, 2015; date of current version September 7, 2015. This work was partially supported by the National Research Fund, Luxembourg, under the projects “CO²SAT”, “SATSENT”, and “SeMiGod”, and by the European Space Agency under “PreDem”. Part of this work has been presented at the IEEE GlobeCom 2014 conference. The associate editor coordinating the review of this paper and approving it for publication was Z. Wang.

The authors are with the Security and Trust (SnT), University of Luxembourg, Luxembourg City L-1359, Luxembourg (e-mail: dimitrios.christopoulos@uni.lu; symeon.chatzinotas@uni.lu; bjorn.ottersten@uni.lu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TWC.2015.2424961

as the most promising use cases of full frequency reuse configurations. Nevertheless, the incorporation of linear precoding techniques is inhibited by the inherent characteristics of the satellite system [1], [2]. The present contribution focuses on two fundamental constraints stemming from the practical system implementation. Firstly, the framing structure of satcom standards, such as the second generation digital video broadcasting for satellite standard DVB-S2 [3] and its most recent extensions DVB-S2X [4], inhibit scheduling a single user per transmission. Secondly, non-flexible on-board payloads prevent power sharing between beams.

Focusing on the first practical constraint, the physical layer design of DVB-S2 [3] has been optimized to cope with the noise limited, with excessive propagation delays and intense fading phenomena, satellite channel. Therefore, long forward error correction (FEC) codes and fade mitigation techniques that rely on an adaptive link layer design (adaptive coding and modulation—ACM) have been employed. The latest evolution of DVB-S2X, through its—synchronous over the multiple beams—superframes (cf. annex E of [4]), allows for the incorporation of the aforementioned interference mitigation techniques (cf. annex C of [5]). A small-scale example of the application of linear precoding methods within the DVB-S2X standard is depicted in Fig. 1. Clearly, the underlying framing structure hinders the calculation of a precoding matrix on a user-by-user basis. During one transmission period, one frame per beam accommodates a different number of users, each with different data requirements. Added to that, the application of FEC block coding over the entire frame requires that co-scheduled users decode the entire frame and then extract the data they need. Also, the unequal data payloads amongst users simultaneously served in different beams further complicates the joint processing of the multiple streams. Consequently, despite the capacity achieving channel based precoding [6], practical system implementations emanate the consideration of precoding on a frame-by-frame basis. The notion of frame-based precoding is presented in more detail in [1], [2].

From a signal processing perspective, physical layer (PHY) multicasting to multiple co-channel groups [7] can provide the theoretically optimal precoders when a multi-antenna transmitter conveys independent sets of common data to distinct groups of users. This scenario is known as PHY multigroup multicast beamforming (or equivalently precoding). The optimality of the multicast multigroup precoders for frame-based precoding is intuitively clear, under the following considerations. In multicasting, the same symbol is transmitted to multiple receivers. This is the fundamental assumption of frame-based precoding

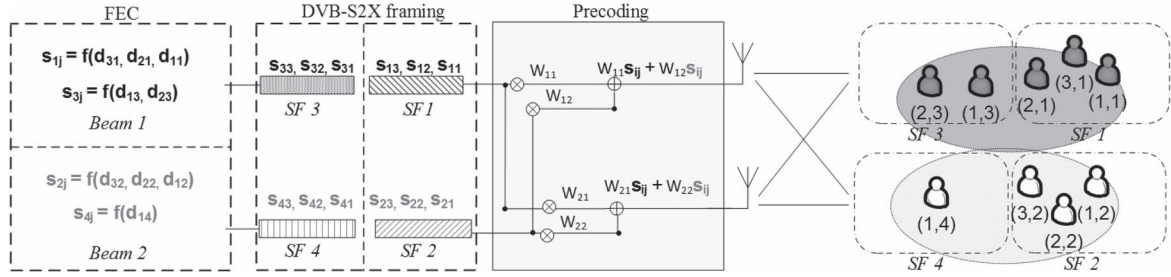


Fig. 1. Frame-based precoding in DVB-S2X. Function $f(\cdot)$ denotes the FEC coding operation over the data d_{xy} that are uniquely addressed to user x of beam y , as identified in the right side of the plot. Consequently, the j -th transmitted symbol s_{ij} , belonging to the i -th superframe (SF), contains an encoded bit-stream that needs to be received by all co-scheduled users. In SFs 3 and 4, different number of users are co-scheduled.

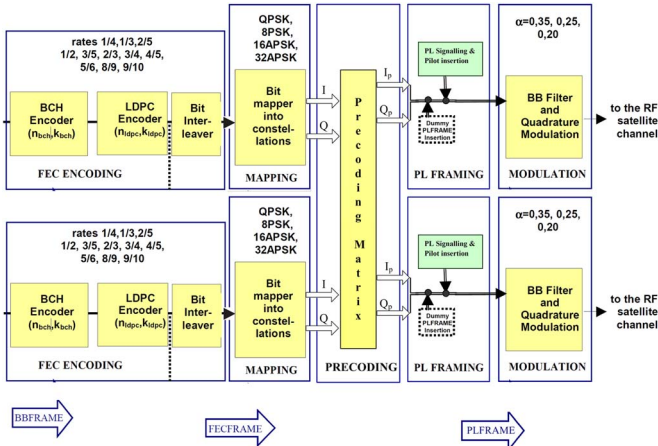


Fig. 2. Transmitter functional block diagram, based on DVB-S2 [3], extended to incorporate advanced interference mitigation techniques.

as well, since the symbols of one frame, regardless of the information they convey, are addressed to multiple users. These users need to receive the entire frame, decode it and then extract information that is relevant to them. The connection between PHY multigroup multicast beamforming (precoding) and frame-based precoding was firstly established in [8].

The second practical constraint tackled in the present work includes a maximum limit on the per-antenna transmitted power. Individual per-antenna amplifiers prevent power sharing amongst the antennas of the future full frequency reuse compatible satellites. On board flexible amplifiers, such as multi-port amplifiers and flexible traveling wave tube amplifiers [9], come at high costs. Also, power sharing is impossible in distributed antenna systems (DAS), such as constellations of cooperative satellite systems (e.g., dual satellite systems [10] or swarms of nano-satellites).

Enabled by the incorporation of linear precoding in DVB-S2X, an example of a full frequency reuse transmission chain is depicted in Fig. 2. The optimal, in a throughput maximizing sense, precoding matrix, combined with a low complexity user scheduling algorithm will be presented in the remaining parts of this work.

A. Related Work

In the PHY multigroup multicast precoding literature, two fundamental optimization criteria, namely the sum power minimization under specific Quality of Service (QoS) constraints

and the maximization of the minimum SINR (max min fair criterion) have been considered in [7], [11], [12] under a SPC. Extending these works, a consolidated solution for the weighted max min fair multigroup multicast beamforming under PACs has been derived in [13], [14]. To this end, the well established tools of Semi-Definite Relaxation (SDR) and Gaussian randomization were combined with bisection to obtain highly accurate and efficient solutions.

The fundamental attribute of multicasting, that is a single transmission to be addressed to a group of users, constrains the system performance according to the worst user. Therefore, the maximization of the minimum SINR is the most relevant problem and the fairness criterion is imperative [13]. When advancing to multigroup multicast systems, however, the service levels between different groups can be adjusted towards achieving some other optimization goal. The sum rate maximization (max SR) problem in the multigroup multicast context was initially considered in [15] under SPC. Therein, a heuristic iterative algorithm based on the principle of decoupling the beamforming design and the power allocation problem was proposed. In more detail, the SPC max sum rate problem was solved using a two step optimization algorithm. The first step was based on the QoS multicast beamforming problem of [7], as iteratively solved with input QoS targets defined by the worst user per group in the previous iteration. The derived precoders push all the users of the group closer to the worst user thus saving power. The second step of the algorithm consisted of the gradient based power reallocation methods of [16]. Hence, a power redistribution takes place via the sub-gradient method [16] to the end of maximizing the system sum rate.

In a realistic system design, the need to schedule a large number of users, over subsequent in time transmissions, is of substantial importance. In the context of multiuser multiple input multiple output (MU-MIMO) communications, user scheduling has shown great potential in maximizing the system throughput performance. In [17], [18], low complexity user scheduling algorithms allowed for the channel capacity approaching performance of linear precoding methods when the number of available users grows large. The enabler for these algorithms is the exact knowledge of the CSI. Motivated by these results and acknowledging that the large number of users served by one satellite can offer significant multiuser diversity gains, channel based user scheduling over satellite is herein proposed. Further supporting this claim, the diverse multiuser satellite environment was exploited towards approaching the information

theoretic channel capacity bounds in [10]. Therein, user scheduling methods were extended to account for adjacent transmitters and applied in a multibeam satellite scenario, exhibiting the importance of scheduling for satcoms. In the present work, drawing intuitions from the frame-based design, multicast-aware user scheduling algorithms are derived. These algorithms, as it will be shown, exploit the readily available CSI, to glean the multiuser diversity gains of satellite systems.

Different from the aforementioned works, the sum rate maximization under PACs has only been considered in [19]. Herein, this principle is used as a stepping stone for the incremental development of elaborate optimization algorithms that solve problems inspired by the needs of frame-based precoding over satellite. The contributions are summarized in the following points:

- The max SR multigroup multicast problem under PACs is formulated and solved.
- The above max SR problem is extended to account for minimum rate constraints (MRCs).
- A novel modulation aware max SR optimization that considers the discretized throughput function of the receive useful signal power is proposed and heuristically solved.
- A low complexity, CSI based, user scheduling algorithm that considers the multigroup multicast nature of the frame-based precoding system is envisaged.
- The developed techniques are evaluated over a multibeam, full frequency reuse satellite scenario.

The rest of the paper is structured as follows. Section II models the multigroup multicast system. Based on this model, the max SR, multigroup multicast optimization problem is formulated and solved in Section III. Extending this optimization, system dependent problems are tackled in Section IV. Further on, user scheduling is discussed in Section V. Finally, in Section VI, the performance of the derived algorithms is evaluated, while Section VII concludes the paper.

Notation: In the remainder of this paper, bold face lower case and upper case characters denote column vectors and matrices, respectively. The operators $(\cdot)^T$, $(\cdot)^\dagger$, $|\cdot|$, $\text{Tr}(\cdot)$, and $\|\cdot\|_2$, correspond to the transpose, the conjugate transpose, the absolute value, the trace and the Euclidean norm operations, while $[\cdot]_{ij}$ denotes the i, j -th element of a matrix. An x -element column vector of ones is denoted as $\mathbf{1}_x$. Finally, \emptyset denotes an empty set.

II. SYSTEM MODEL

The focus is on a single broadband multibeam satellite transmitting to multiple single antenna users. Let N_t denote the number of transmitting elements, which for the purposes of the present work, are considered equal to the number of beams (one feed per beam assumption) and N_u the total number of users simultaneously served. The received signal at the i -th user will read as $y_i = \mathbf{h}_i^\dagger \mathbf{x} + n_i$, where \mathbf{h}_i^\dagger is a $1 \times N_t$ vector composed of the channel coefficients (i.e. channel gains and phases) between the i -th user and the N_t antennas of the transmitter, \mathbf{x} is the $N_t \times 1$ vector of the transmitted symbols and n_i is the complex circular symmetric (c.c.s.) independent identically distributed (i.i.d) zero mean Additive White Gaussian Noise

(AWGN), measured at the i -th user's receiver. Herein, for simplicity, the noise will be normalized to one and the impact of noise at the receiver side will be incorporated in the channel coefficients, as will be shown in the following [Section II-A (4)].

Let us assume that a total of N_t multicast groups are realized where $\mathcal{I} = \{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_{N_t}\}$ the collection of index sets and \mathcal{G}_k the set of users that belong to the k -th multicast group, $k \in \{1 \dots N_t\}$. Each user belongs to only one frame (i.e., group), thus $\mathcal{G}_i \cap \mathcal{G}_j = \emptyset$, $\forall i, j \in \{1 \dots N_t\}$, while $\rho = N_u/N_t$ denotes the number of users per group. Let $\mathbf{w}_k \in \mathbb{C}^{N_t \times 1}$ denote the precoding weight vector applied to the transmit antennas to beamform towards the k -th group of users. By collecting all user channels in one channel matrix, the general linear signal model in vector form reads as $\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} = \mathbf{H}\mathbf{W}\mathbf{s} + \mathbf{n}$, where \mathbf{y} and $\mathbf{n} \in \mathbb{C}^{N_u}$, $\mathbf{x} \in \mathbb{C}^{N_t}$ and $\mathbf{H} \in \mathbb{C}^{N_u \times N_t}$. Since, the frame-based precoding imposes a single precoding vector for multiple users, the matrix will include as many precoding vectors (i.e., columns) as the number of multicast groups. This is the number of transmit antennas, since one frame per-antenna is assumed. Also, the symbol vector includes a single equivalent symbol for each frame i.e., $\mathbf{s} \in \mathbb{C}^{N_t}$, inline with the multicast assumptions. Consequently, a square precoding matrix is realized, i.e., $\mathbf{W} \in \mathbb{C}^{N_t \times N_t}$. The assumption of independent information transmitted to different frames implies that the symbol streams $\{s_k\}_{k=1}^{N_t}$ are mutually uncorrelated. Also, the average power of the transmitted symbols is assumed normalized to one. Therefore, the total power radiated from the antenna array is equal to

$$P_{tot} = \sum_{k=1}^{N_t} \mathbf{w}_k^\dagger \mathbf{w}_k = \text{Trace}(\mathbf{W}\mathbf{W}^\dagger), \quad (1)$$

where $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{N_t}]$. The power radiated by each antenna element is a linear combination of all precoders and reads as [20]

$$P_n = \left[\sum_{k=1}^{N_t} \mathbf{w}_k \mathbf{w}_k^\dagger \right]_{nn} = [\mathbf{W}\mathbf{W}^\dagger]_{nn}, \quad (2)$$

where $n \in \{1 \dots N_t\}$ is the antenna index. The fundamental difference between the SPC of [7] and the proposed PAC is clear in (2), where instead of one, N_t constraints are realized, each one involving all the precoding vectors.

A. Multibeam Satellite Channel

The above general system model is applied over a multibeam satellite channel explicitly defined as follows. A 245 beam pattern that covers Europe is employed [22]. For the purposes of the present work, only a subset of the 245 beams will be considered, as presented in Fig. 3. Such a consideration is in line with the multiple gate-way (multi-GW) assumptions of large multibeam systems [21]. However, the effects of interference from adjacent clusters is left for future investigations. A complex channel matrix that models the link budget of each user as well as the phase rotations induced by the signal propagation is employed in the standards of [8], [9], and [22]. In more detail, the total channel matrix $\mathbf{H} \in \mathbb{C}^{N_u \times N_t}$ is generated as

$$\mathbf{H} = \Phi \mathbf{B}, \quad (3)$$

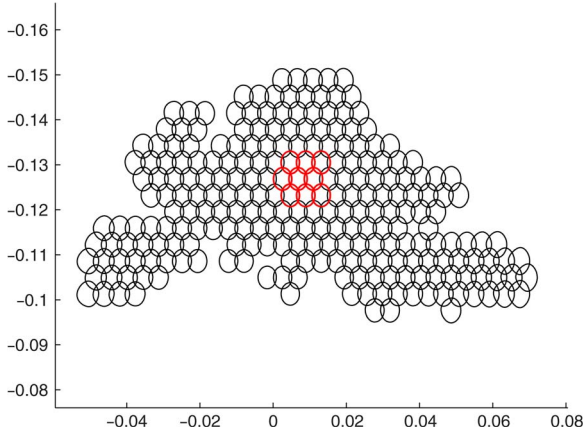


Fig. 3. Beam pattern covering Europe, provided by [22], with the nine beams considered herein highlighted.

and includes the multibeam antenna pattern (matrix \mathbf{B}) and the signal phase due to different propagation paths between the users (matrix Φ). The real matrix $\mathbf{B} \in \mathbb{R}^{N_u \times N_t}$ models the satellite antenna radiation pattern, the path loss, the receive antenna gain and the noise power. Its i, j -th entry is given by [22]:

$$b_{ij} = \left(\frac{\sqrt{G_R G_{ij}}}{4\pi(d_k \cdot \lambda^{-1})\sqrt{\kappa T_{cs} B_u}} \right), \quad (4)$$

with d_k the distance between the i -th user and the satellite (slant-range), λ the wavelength, κ the Boltzman constant, T_{cs} the clear sky noise temperature of the receiver, B_u the user link bandwidth, G_R the receiver antenna gain and G_{ij} the multibeam antenna gain between the i -th single antenna user and the j -th on board antenna (= feed). Hence, the beam gain for each satellite antenna-user pair, depends on the antenna pattern and on the user position.

An inherent characteristic of the multibeam satellite channel is the high correlation of signals at the satellite side. Thus a common assumption in multibeam channel models is that each user will have the same phase between all transmit antennas due to the long propagation path [9]. The identical phase assumption between one user and all transmit feeds is supported by the relatively small distances between the transmit antennas and the long propagation distance of all signals to a specific receiver. Hence, in (3) the diagonal square matrix Φ is generated as $[\Phi]_{xx} = e^{j\phi_x}$, $\forall x = 1 \dots N_u$ where ϕ_x is a uniform random variable in $[2\pi, 0)$ and $[\Phi]_{xy} = 0$, $\forall x \neq y$.

B. Average User Throughput

Based on the above link budget considerations, the achievable average user throughput is normalized over the number of beams, to provide a metric comparable with multibeam systems of any size. Therefore, the average user throughput, R_{avg} as will be hereafter referred to, is given as

$$R_{avg} = \frac{2B_u}{1 + \alpha} \frac{1}{N_t} \sum_{k=1}^{N_t} f_{\text{DVB-S2X}} \left(\min_{i \in \mathcal{G}_k} \{\text{SINR}_i\}, \mathbf{t} \right), \quad (5)$$

in [Gbps/beam], where all parameters are defined in Table II of Section VI. In (5), the spectral efficiency function $f_{\text{DVB-S2X}}$ receives as input each users SINR as well as a threshold vector \mathbf{t} . Then, $f_{\text{DVB-S2X}}$ performs a rounding of the input SINR to the closest lower floor given by the threshold vector \mathbf{t} and outputs the corresponding spectral efficiency in [bps/Hz]. This operation is denoted as $\lfloor \cdot \rfloor_{\mathbf{t}}$. The mapping of receive SINR regions to a spectral efficiency achieved by a respective modulation and coding (MODCOD) scheme is explicitly defined in the latest evolution of the satcom standards [4]. It should also be noted, that the conventional four color frequency reuse calculations are based on the exact same formula, with the only modifications being the input SINR, calculated under conventional four color reuse pattern and with the pre-log factor reduced by four times, equal to the conventional fractional frequency reuse [22].

III. SUM RATE MAXIMIZATION

For the precoding design, optimal multigroup multicast precoders under per-antenna constraints are proposed to maximize the throughput of the multibeam satellite system. The design of throughput maximizing optimal precoders is a complicated problem without an explicit solution even for the unicasting case [23]. When advancing to multicasting assumptions, the structure of the problem becomes even more involved, as already explained [11]. Consequently, the present work builds upon the heuristic methods of [15], [16].

Since a multigroup multicasting scenario entails the flexibility to maximize the total system rate by providing different service levels amongst groups, the multigroup multicast max SR optimization aims at increasing the minimum SINR within each group while in parallel maximizing the sum of the rates of all groups. Intuitively, this can be accomplished by reducing the SINR of users with better conditions than the worst user of their group. Also, groups that contain compromised users might need to be turned off, hence driving their users to service unavailability, to save power resources and degrees of freedom. As a result, power is not consumed for the mitigation of poor channel conditions. Any remaining power budget is then reallocated to well conditioned and balanced in terms of performance groups.

A. Per-Antenna Power Constrained Optimization

This section focuses on the per-antenna power constrained max SR problem, formally defined as

$$\begin{aligned} \mathcal{SR} : \quad & \max_{\{\mathbf{w}_k\}_{k=1}^{N_t}} \sum_{i=1}^{N_u} \log_2(1 + \gamma_i) \\ \text{subject to :} \quad & \gamma_i = \min_{m \in \mathcal{G}_k} \frac{|\mathbf{w}_k^\dagger \mathbf{h}_m|^2}{\sum_{l \neq k} |\mathbf{w}_l^\dagger \mathbf{h}_m|^2 + \sigma_m^2}, \\ & \forall i \in \mathcal{G}_k, k, l \in \{1 \dots N_t\}, \\ \text{and to :} \quad & \left[\sum_{k=1}^{N_t} \mathbf{w}_k \mathbf{w}_k^\dagger \right]_{nn} \leq P_n, \\ & \forall n \in \{1 \dots N_t\}. \end{aligned} \quad (6)$$

Problem \mathcal{SR} receives as input the channel matrices as well as the per-antenna power constraint vector $\mathbf{p}_{ant} = [P_1, P_2, \dots, P_{N_t}]$. Following the notation of [7] for ease of reference, the optimal objective value of \mathcal{SR} will be denoted as $c^* = \mathcal{SR}(\mathbf{p}_{ant})$.

and the associated optimal point as $\{\mathbf{w}_k^{\mathcal{SR}}\}_{k=1}^{N_t}$. The novelty of the \mathcal{SR} lies in the PACs, i.e., (7) instead of the conventional SPC proposed in [15]. Therein, to solve the elaborate max SR under a SPC problem, the decoupling of the precoder calculation and the power loading over these vectors was considered. The first problem was solved based on the solutions of [7] while the latter on sub-gradient optimization methods [16]. To the end of solving the novel \mathcal{SR} problem, a heuristic algorithm is proposed herein. Different than in [15], the new algorithm calculates the per-antenna power constrained precoders by utilizing recent results [13]. Also, modified sub-gradient optimization methods are proposed to take into account the PACs. More specifically, instead of solving the QoS sum power minimization problem of [7], the proposed algorithm calculates the PAC precoding vectors by solving the following problem [13] that reads as

$$\begin{aligned} \mathcal{Q} : \quad & \min_{r, \{\mathbf{w}_k\}_{k=1}^{N_t}} r \\ & \text{subject to : } \frac{|\mathbf{w}_k^\dagger \mathbf{h}_i|^2}{\sum_{l \neq k} |\mathbf{w}_l^\dagger \mathbf{h}_i|^2 + \sigma_i^2} \geq \gamma_i, \\ & \quad \forall i \in \mathcal{G}_k, k, l \in \{1 \dots N_t\}, \\ & \text{and to : } \frac{1}{P_n} \left[\sum_{k=1}^{N_t} \mathbf{w}_k \mathbf{w}_k^\dagger \right]_{nn} \leq r, \\ & \quad \forall n \in \{1 \dots N_t\}, \end{aligned} \quad (8)$$

$$\text{and to : } \frac{1}{P_n} \left[\sum_{k=1}^{N_t} \mathbf{w}_k \mathbf{w}_k^\dagger \right]_{nn} \leq r, \quad (9)$$

where $r \in \mathbb{R}^+$. Problem \mathcal{Q} receives as input the SINR target vector $\mathbf{g} = [\gamma_1, \gamma_2, \dots, \gamma_{N_u}]$, that is the individual QoS constraints of each user, as well as the per-antenna power constraint vector \mathbf{p}_{ant} . Let the optimal objective value of \mathcal{Q} be denoted as $r^* = \mathcal{Q}(\mathbf{g}, \mathbf{p}_{ant})$ and the associated optimal point as $\{\mathbf{w}_k^{\mathcal{Q}}\}_{k=1}^{N_t}$. This problem is solved using the well established methods of SDR and Gaussian randomization [24]. A more detailed description of the solution of \mathcal{Q} can be found in [13], [14] and is herein omitted for conciseness.

To proceed with the power reallocation step, let us rewrite the precoding vectors calculated from \mathcal{Q} as $\{\mathbf{w}_k^{\mathcal{Q}}\}_{k=1}^{N_t} = \{\sqrt{p_k} \mathbf{v}_k\}_{k=1}^{N_t}$ with $\|\mathbf{v}_k\|_2^2 = 1$ and $\mathbf{p} = [p_1 \dots p_k]$. By this normalization, the beamforming problem can be decoupled into two problems. The calculation of the beamforming directions, i.e., the normalized $\{\mathbf{v}_k\}_{k=1}^{N_t}$, and the power allocation over the existing groups, i.e., the calculation of \mathbf{p}_k . Since the exact solution of \mathcal{SR} is not straightforwardly obtained, this decoupling allows for a two step optimization. Under general unicasting assumptions, the SR maximizing power allocation with fixed beamforming directions is a convex optimization problem [16]. Nonetheless, when multigroup multicasting is considered, the cost function $C_{\mathcal{SR}} = \sum_{k=1}^{N_t} \log(1 + \min_{i \in \mathcal{G}_k} \{\text{SINR}_i\})$ is no longer differentiable due to the $\min_{i \in \mathcal{G}_k}$ operation and one has to adhere to sub-gradient solutions [15]. What is more, as in detail explained in [15], the cost function needs to be continuously differentiable, strictly increasing, with a log-convex inverse function. Nevertheless, this is not the case for \mathcal{SR} . Towards providing a heuristic solution to an involved problem without known optimal solution, an optimization over the logarithmic power vector $\mathbf{s} = \{s_k\}_{k=1}^{N_t} = \{\log p_k\}_{k=1}^{N_t}$, will be considered in the standards of [15]. Therein, the authors employ a function ϕ that satisfies the above assumptions to approximate the utility

function of \mathcal{SR} . For more information on function ϕ and the suggested approximation, the reader is directed to [15]. It should be noted that the heuristic nature of this solution does not necessarily guarantee convergence to a global optimum. Albeit this, and despite being sub-optimal in the max sum rate sense, the heuristic solutions attain a good performance, as shown in [15], [16] and in the following. Consequently, in the present contribution, the power loading is achieved via the sub-gradient method [16], under specific modifications over [15] that are hereafter described.

The proposed algorithm, presented in Algorithm 1, is an iterative two step procedure. In each step, the QoS targets \mathbf{g} are calculated as the minimum target per group of the previous iteration, i.e., $\gamma_i = \min_{i \in \mathcal{G}_k} \{\text{SINR}_i\}$, $\forall i \in \mathcal{G}_k, k \in \{1 \dots N_t\}$. Therefore, the new precoders require equal or less power to achieve the same system sum rate. Any remaining power is then redistributed amongst the groups to the end of maximizing the total system throughput, via the sub-gradient method [16]. Focusing of the later method and using the logarithmic power vector $\mathbf{s} = \{s_k\}_{k=1}^{N_t} = \{\log p_k\}_{k=1}^{N_t}$, the sub-gradient search method is given as

$$\mathbf{s}(t+1) = \prod_{\mathbb{P}} [\mathbf{s}(t) - \delta(t) \cdot \mathbf{r}(t)], \quad (10)$$

where $\prod_{\mathbb{P}}[\mathbf{x}]$ denotes the projection operation of point $\mathbf{x} \in \mathbb{R}_{N_t}$ onto the set $\mathbb{P} \subset \mathbb{R}_{N_t}^+$. The parameters $\delta(t)$ and $\mathbf{r}(t)$ are the step of the search and the sub-gradient of the \mathcal{SR} cost function at the point $\mathbf{s}(t)$, respectively. The number of iterations this method runs, denoted as t_{\max} , is predefined. The projection operation, i.e., $\prod_{\mathbb{P}}[\cdot]$, constrains each iteration of the sub-gradient to the feasibility set of the \mathcal{SR} problem. The analytic calculation of $\mathbf{r}(t)$ follows the exact steps of [15], [16] and is herein omitted for shortness. To account for the more complicated PACs the projection over a per-antenna power constrained set is considered as follows. The set of PACs can be defined as

$$\mathbb{P} = \left\{ \mathbf{p} \in \mathbb{R}_{N_t}^+ \mid \left[\sum_{k=1}^{N_t} p_k \mathbf{v}_k \mathbf{v}_k^\dagger \right]_{nn} \leq P_n \right\}, \quad (11)$$

where the elements of the power vector $\mathbf{p} = \exp(\mathbf{s})$ represent the power allocated to each group. It should be stressed that this power is inherently different from the power transmitted by each antenna $\mathbf{p}_{ant} \in \mathbb{R}_{N_t}^+$. The connection between \mathbf{p}_{ant} and \mathbf{p} is given by the normalized beamforming vectors as easily observed in (11). Different from the sum power constrained solutions of [15], the per-antenna constrained projection problem is given by

$$\begin{aligned} \mathcal{P} : \quad & \min_{\mathbf{p}} \|\mathbf{p} - \mathbf{x}\|_2^2 \\ & \text{subject to : } \left[\sum_{k=1}^{N_t} p_k \mathbf{v}_k \mathbf{v}_k^\dagger \right]_{nn} \leq P_n, \\ & \quad \forall n \in \{1 \dots N_t\}, \end{aligned} \quad (12)$$

where $\mathbf{p} \in \mathbb{R}_{N_t}$ and $\mathbf{x} = \exp(\mathbf{s}(t))$. Problem \mathcal{P} is a quadratic problem (QP) [25] and can thus be solved to arbitrary accuracy using standard numerical methods.¹ Subsequently, the solution

¹ Analytical methods to solve problem \mathcal{P} are beyond the scope of the present work. For more information, the reader is referred to [25].

TABLE I
INPUT PARAMETERS FOR ALGORITHM 1

Parameter	Symbol	Value
Sub-gradient iterations	t_{max}	1
Sub-gradient initial value	$\delta(t)$	0.4
Sub-gradient step	$\delta(t+1)$	$\delta(t)/2$
Gaussian Randomizations	N_{rand}	100
Per-antenna constraints	\mathbf{p}_{ant}	$P_{tot}/N_t \cdot \mathbf{1}_{N_t}$
User Noise variance	σ_i^2	1, $\forall i \in \{1 \dots N_u\}$

of (10) is given as $\mathbf{s}(l+1) = \log(\mathbf{p}^*)$, where $\mathbf{p}^* = \mathcal{P}(\mathbf{p}_{ant}, \mathbf{x})$ is the optimal point of convex problem \mathcal{P} . To summarize the solution process, the per-antenna power constrained sum rate maximizing algorithm is given in Algorithm 1.

Algorithm 1 Sum-rate maximizing multigroup multicasting under per-antenna power constraints.

Input: (see Table I) $\{\mathbf{w}_k^{(0)}\}_{k=1}^{N_t} = \sqrt{P_{tot}/(N_t^2)} \cdot \mathbf{1}_{N_t}$, \mathbf{p}_{ant} , $j = 0$.

Output: $\{\mathbf{w}_k^{SR}\}_{k=1}^{N_t}$

begin

while SR does not converge **do**

$j = j + 1$

Step 1: Solve $\mathbf{r}^* = Q(\mathbf{g}_{(j)}, \mathbf{p}_{ant})$ to calculate $\{\mathbf{w}_k^{(j)}\}_{k=1}^{N_t}$. The input SINR targets $\mathbf{g}_{(j)}$ are given by the minimum SINR per group, i.e., $\gamma_i = \min_{i \in \mathcal{G}_k} \{\text{SINR}_i\}$, $\forall i \in \mathcal{G}_k$, $k \in \{1 \dots N_t\}$.

Step 2: Initialize the sub-gradient search algorithm as: $\mathbf{p}^{(j)} = \{p_k\}_{k=1}^{N_t} = \{\|\mathbf{w}_k^{(j)}\|_2^2\}_{k=1}^{N_t}$, $\mathbf{s}^{(j)} = \{s_k\}_{k=1}^{N_t} = \{\log p_k\}_{k=1}^{N_t}$, $\{\mathbf{v}_k^{(j)}\}_{k=1}^{N_t} = \{\mathbf{w}_k^{(j)} / \sqrt{p_k^{(j)}}\}_{k=1}^{N_t}$.

Step 3: Calculate t_{max} iterations of the sub-gradient power control algorithm, starting from $\mathbf{s}(0) = \mathbf{s}^{(j)}$:

for $t = 0 \dots t_{max} - 1$ **do**

$\mathbf{s}(t+1) = \prod_{\mathbb{P}} [\mathbf{s}(t) - \delta(t) \cdot \mathbf{r}(t)]$

end

$\mathbf{s}^{(j+1)} = \mathbf{s}(t_{max} - 1)$,

Step 4: Calculate the current throughput: $c^* = SR(\mathbf{p}_{ant})$ with $\{\mathbf{w}_k^{SR}\}_{k=1}^{N_t} = \{\mathbf{w}_k^{(j+1)}\}_{k=1}^{N_t} = \{\mathbf{v}_k^{(j)} \exp(s_k^{(j+1)})\}_{k=1}^{N_t}$

end

B. Complexity and Convergence Analysis

An important discussion involves the complexity of the proposed algorithm. In [13], [14], the computational burden for an accurate approximate solution of the per-antenna power minimization problem \mathcal{Q} (step 1 of Alg. 1) has been calculated. In summary, the relaxed power minimization is an semidefinite programming (SDP) instance with N_t matrix variables of $N_t \times N_t$ dimensions and $N_u + N_t$ linear constraints. The present work relies on the CVX tool [25] which calls numerical solvers such as SeDuMi to solve semi-definite programs. The interior point methods employed to solve this SDP require at most $\mathcal{O}(N_t \log(1/\epsilon))$ iterations, where ϵ is the desired numerical accuracy of the solver. Moreover, in each iteration not more than $\mathcal{O}(N_t^9 + N_t^4 + N_u N_t^3)$ arithmetic operations will be performed.

The solver used [25] also exploits the specific structure of matrices hence the actual running time is reduced. Next, a fixed number of iterations of the Gaussian randomization method is performed [24]. In each randomization, a linear problem (LP) is solved with a worst case complexity of $\mathcal{O}(N_t^{3.5} \log(1/\epsilon_1))$ for an ϵ_1 -optimal solution. The accuracy of the solution increases with the number of randomizations [7], [11], [24]. The remaining three steps of Alg. 1 involve a closed form sub-gradient calculation as given in [16] and the projection operation, which is a real valued least square problem under N_t quadratic inequality PACs. Consequently, the asymptotic complexity of the derived algorithm is polynomial, dominated by the complexity of the QoS multigroup multicast problem under PACs.

The convergence of Alg. 1 is guaranteed given that the chosen step size satisfies the conditions given in [15], [16], that is the diminishing step size. Herein, $\delta(l+1) = \delta(l)/2$. What is more, in accordance to [15], only a single iteration of the sub-gradient is performed in the numerical results (i.e. $t_{max} = 1$).

IV. SYSTEM DRIVEN OPTIMIZATION

Constraints inspired by the inherent nature of satellite communications emanate the definition of novel optimization problems. The present section focuses on enabling demanding in terms of availability satellite services. Increased scepticism over spectrally efficient, aggressive frequency reuse, multibeam satellites stems from the effects of such configurations on the SINR distribution across the coverage. In full frequency reuse scenarios, the useful signal power at the receiver is greatly reduced due to the intra-system interferences. Despite the throughput gains due to the increased user link bandwidth and the adequate management of interferences by linear precoding, the mean and variance of the SINR distribution over the coverage area is generally reduced. This is the price paid for increasing the frequency reuse. Naturally, this reduction in the average SINR will lead to a higher utilization of lower MODCODs and increase the probability of service unavailability over the coverage (outage probability). Retransmissions that incur in these outage instances, are bound to burden the system in terms of efficiency. What is more, by acknowledging the multiuser satellite environment (cf. Section V), these outage periods can potentially become comparable to the inherent long propagation delay of satcoms. Such a case will render the overall delay, as experienced by a user, unacceptable. As a result, the probability of compromised users to experience long outage periods, needs to be considered in a system level. In this work, the introduction of minimum rate constraints over the entire coverage is proposed, as a means to guarantee in the physical layer design the stringent availability requirements typically accustomed in satcoms. The guarantee of a minimum level of service availability is introduced for the first time in a max SR multigroup multicast optimization.

A. Sum Rate Maximization Under Minimum Rate Constraints

To provide high service availability, the gains of the sum rate optimization can be traded-off in favor of a minimum guaranteed rate across the coverage. This trade-off mostly

depends on the minimum MODCOD supported by the ACM.² Since an intermediate solution between the fairness and the max SR goals is of high engineering interest, a novel optimization problem, namely the throughput maximization under availability constraints, is proposed. The innovation, aspired by operational requirements, lies in the incorporation of minimum rate constraints (MRCs) in the PAC sum rate maximizing problem (equivalently minimum SINR constraints). Formally, the new optimization problem is defined as

$$\begin{aligned} \mathcal{SRA} : \quad & \max_{\{\mathbf{w}_k\}_{k=1}^{N_t}} \sum_{i=1}^{N_u} \log_2(1 + \gamma_i) \\ \text{s.t. : } & \gamma_i = \min_{m \in \mathcal{G}_k} \frac{|\mathbf{w}_k^\dagger \mathbf{h}_m|^2}{\sum_{l \neq k} |\mathbf{w}_l^\dagger \mathbf{h}_m|^2 + \sigma_m^2}, \\ & \forall i \in \mathcal{G}_k, k, l \in \{1 \dots N_t\}, \\ \text{and to : } & \left[\sum_{k=1}^{N_t} \mathbf{w}_k \mathbf{w}_k^\dagger \right]_{nn} \leq P_n, \\ & \forall n \in \{1 \dots N_t\}, \\ \text{and to : } & \gamma_i \geq \gamma_{\min}, \forall i \in \{1 \dots N_u\}. \end{aligned} \quad (13)$$

$$\text{and to : } \left[\sum_{k=1}^{N_t} \mathbf{w}_k \mathbf{w}_k^\dagger \right]_{nn} \leq P_n, \quad (14)$$

$$\text{and to : } \gamma_i \geq \gamma_{\min}, \forall i \in \{1 \dots N_u\}. \quad (15)$$

In \mathcal{SRA} , the power allocation needs to account for the MRCs, i.e., (15). This is achieved by modifying the constraints of the sub-gradient search [16], as imposed via the projection of the current power vector onto the convex set of constraints. Therefore, the additional constraint can be introduced in the projection method, since it does not affect the convexity of the formulation. Subsequently, to solve \mathcal{SRA} a new projection that includes the minimum rate constraints is proposed. The new subset, that is the min SINR constrained set, is a convex subset of the initially convex set. The availability constrained projection reads as

$$\begin{aligned} \mathcal{PA} : \quad & \min_{\mathbf{p}} \|\mathbf{p} - \mathbf{x}\|_2^2 \\ \text{subject to : } & \frac{p_k |\mathbf{v}_k^\dagger \mathbf{h}_i|^2}{\sum_{l \neq k} p_l |\mathbf{v}_l^\dagger \mathbf{h}_i|^2 + \sigma_i^2} \geq \gamma_{\min}, \\ & \forall i \in \mathcal{G}_k, k, l \in \{1 \dots N_t\}, \\ \text{and to : } & \left[\sum_{k=1}^{N_t} p_k \mathbf{v}_k \mathbf{v}_k^\dagger \right]_{nn} \leq P_n, \\ & \forall n \in \{1 \dots N_t\}, \end{aligned} \quad (16)$$

$$\text{and to : } \left[\sum_{k=1}^{N_t} p_k \mathbf{v}_k \mathbf{v}_k^\dagger \right]_{nn} \leq P_n, \quad (17)$$

which is a convex optimization problem, that includes one additional linear constraint, i.e., (16), over \mathcal{P} . Provided that \mathcal{SRA} is feasible, then (15) is satisfied and thus a solution for \mathcal{PA} always exists. Similarly to \mathcal{P} , this problem can be solved using standard methods [25].

Subsequently, the solution of \mathcal{SRA} is derived following the steps of Algorithm 1 but with a modification in the sub-gradient method (Step 3), where the projection is calculated by solving problem \mathcal{PA} instead of \mathcal{P} . As intuitively expected, the introduction of MRCs is bound to decrease the system throughput

performance. However, this trade-off can be leveraged towards more favorable conditions, by considering other system aspects, as will be discussed in the following.

B. Throughput Maximization via MODCOD Awareness

A modulation constrained practical system employs higher order modulations to increase its rate with respect to the useful signal power. The strictly increasing logarithmic cost functions describe communications based on Gaussian alphabets and provide the Shannon upper bound of the system spectral efficiency. Therefore, the sum rate maximization problems solved hitherto fail to account for the modulation constrained throughput performance of practical systems. The complication lies in the analytically intractable, at least by the methods considered herein, nature of a step cost function. In the present section, an attempt to leverage this cost function in favor of the system throughput performance is presented. In more detail, benefiting from the finite granularity of the rate function (5) over the achieved SINR, an extra system level optimization can be defined as

$$\begin{aligned} \mathcal{SRM} : \quad & \max_{\{\mathbf{w}_k\}_{k=1}^{N_t}} \sum_{i=1}^{N_u} f_{\text{DVB-S2X}}(\gamma_i, \mathbf{t}) \\ \text{s.t. : } & \gamma_i = \min_{m \in \mathcal{G}_k} \frac{|\mathbf{w}_k^\dagger \mathbf{h}_m|^2}{\sum_{l \neq k} |\mathbf{w}_l^\dagger \mathbf{h}_m|^2 + \sigma_m^2}, \\ & \forall i \in \mathcal{G}_k, k, l \in \{1 \dots N_t\}, \\ \text{and to : } & \left[\sum_{k=1}^{N_t} \mathbf{w}_k \mathbf{w}_k^\dagger \right]_{nn} \leq P_n, \\ & \forall n \in \{1 \dots N_t\}, \\ \text{and to : } & \gamma_i \geq \gamma_{\min}, \forall i \in \{1 \dots N_u\}, \end{aligned} \quad (18)$$

$$\text{and to : } \left[\sum_{k=1}^{N_t} \mathbf{w}_k \mathbf{w}_k^\dagger \right]_{nn} \leq P_n, \quad (19)$$

$$\text{and to : } \gamma_i \geq \gamma_{\min}, \forall i \in \{1 \dots N_u\}, \quad (20)$$

where $f_{\text{DVB-S2X}}(\cdot, \cdot)$ is the finite granularity step function defined in (5). The realization of a non-strictly increasing cost function inhibits the application of gradient based solutions and necessitates a different solution process. To provide a solution for this elaborate -yet of high practical value- problem, a heuristic iterative algorithm is proposed. More specifically, Algorithm 2 receives as input the availability constrained precoders $\{\mathbf{w}_k^{\mathcal{SRA}}\}_{k=1}^{N_t}$ calculated as described in Section IV-A, and calculates an initial SINR distribution. Then, it derives new precoding vectors under minimum SINR constraints given by the closest lower threshold of the worst user in each group, according to the discrete throughput function. Therefore, the resulting system throughput is not decreased while power is saved. This power can now be redistributed. Also, in this manner, the solution guarantees a minimum system availability. Following this step, an ordering of the groups takes place, in terms of minimum required power to increase each group to the next threshold target. For this, the power minimization problem is executed for each group. Next, each of the available groups, starting from the group that requires the least power, is sequentially given a higher target. With the new targets, the power minimization problem is again solved. This constitutes a feasibility optimization check. If the required power satisfies the per antenna constraints, then these precoders are kept. Otherwise the current group is given its previous feasible SINR target and the search proceeds to the next group.

²For instance in DVB-S2X under normal operation over a linearized channel, the most robust modulation and coding rate can provide quasi error free communications (frame error probability lower than 10^{-5}) for as low as -2.85 dB of user SINR, thus achieving a minimum spectral efficiency of 0.4348 [bps/Hz] [5]. Beyond this value, a service outage occurs.

Remark: A further improved solution can be attained when dropping the constraint of a single step increase per group. Herein, such a consideration is avoided for complexity reasons. Since each of the N_t groups can take at most N_m possible SINR values, where N_m denotes the number of MODCODs, by allowing each group to increase more than one step, the number of possible combinations can be as much as $(N_m)^{N_t}$. As a result, the complexity of the optimal solution found by searching the full space of possible solutions, grows exponentially with the number of groups. In the present work, the high number of threshold values for $f_{\text{DVB-S2X}}$ prohibits such considerations.

The summary of this algorithm is given in Algorithm 2. Since it is an iterative algorithm over the number of available groups, convergence is guaranteed. Also, since it receives as input the \mathcal{SRA} solution, its complexity is dominated by the complexity of Algorithm 1, as described in Section III-B.

Algorithm 2 Discretized sum rate maximization.

Input: $\mathbf{H}, P_{\text{tot}}, \sigma_i^2 \quad \forall i \in \{1 \dots N_t\}, \{\mathbf{w}_k^{(0)}\}_{k=1}^{N_t} = \{\mathbf{w}_k^{\mathcal{SRA}}\}_{k=1}^{N_t}, r^{(0)}, \gamma_{\min}$
Output: $\{\mathbf{w}_k^{\text{out}}\}_{k=1}^{N_t}$
begin
 $j = 0; q = 1; \{\mathbf{w}_k^{\text{out}}\}_{k=1}^{N_t} = \{\mathbf{w}_k^{(0)}\}_{k=1}^{N_t};$
Step 1: Solve $r^{*,(0)} = \mathcal{Q}(\mathbf{g}^{(0)}, \mathbf{p}_{\text{ant}})$ to calculate $\{\mathbf{w}_k^{\mathcal{Q},(0)}\}_{k=1}^{N_t}$.
 The input SINR targets are given by the minimum threshold SINR per group, i.e., $\mathbf{g}^{(0)}: \gamma_i = \lfloor \min_{m \in \mathcal{G}_k} \{\text{SINR}_m\} \rfloor_t$, $\forall i, m \in \mathcal{G}_k, k = 1, \dots, N_t$.
for $j = 1 \dots N_t$ **do**
 Step 2: Solve $r^{*,(j)} = \mathcal{Q}(\mathbf{g}^{(j)}, \mathbf{p})$ to calculate $\{\mathbf{w}_k^{\mathcal{Q},(j)}\}_{k=1}^{N_t}$.
 The targets of the current j -th group are increased by one level: $\gamma_i = \lceil \min_{m \in \mathcal{G}_j} \{\text{SINR}_m\} \rceil_t$, $\forall i \in \mathcal{G}_j$;
 Order the groups in terms of increasing $r^{*,(j)}$.
end
while $r^{*,(q)} < 1$ **do**
 Step 3: For each group, in a sequence ordered by the previous step, increase the target by one level;
 Solve $r^{*,(q)} = \mathcal{Q}(\mathbf{g}^{(q)}, \mathbf{p})$ with input targets from the previous iteration: $\mathbf{g}^{(q)} = \mathbf{g}^{(q-1)}$; $q = q + 1$
end
 $\{\mathbf{w}_k^{\text{out}}\}_{k=1}^{N_t} = \{\mathbf{w}_k^{\mathcal{Q},(q)}\}_{k=1}^{N_t}$
end

V. USER SCHEDULING

Multibeam satellite systems typically cover vast areas by a single satellite illuminating a large pool of users requesting service. Therefore, a satcom system operates in a large multiuser environment. In current satcom standards, user scheduling is based on the traffic demand and channel quality [3]. Thus DVB-S2 schedules relatively similar in terms of SINR users in the same frame and a specific link layer mode (assuming ACM) is employed to serve them. A diagram with the necessary operations performed at the transmitter is illustrated in Fig. 4(a) for conventional systems. In aggressive resource reuse transmitters that employ precoding, scheduling policies can be based on the principles of MU-MIMO communications. The inherent

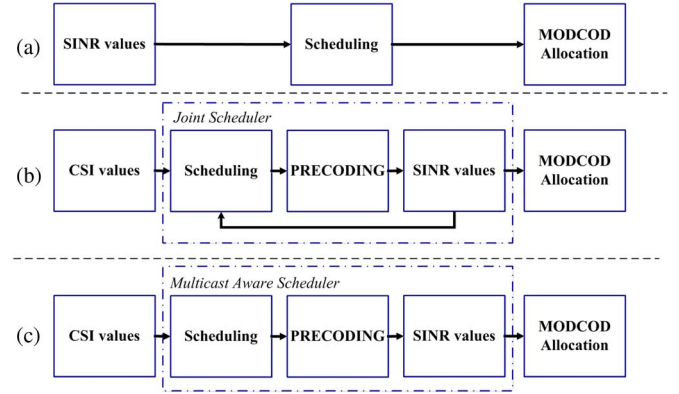


Fig. 4. Scheduling over satellite: (a) Conventional DVB-S2. (b) Optimal joint precoding and scheduling. (c) Proposed multicast-aware heuristic scheduling.

difference with conventional systems is that the CSI for each user is now an N_t dimensional vector rather than a single SINR value. In the parlance of MU-MIMO communications the level of similarity between the users can be measured in terms of orthogonality of the complex vector channels. To maximize the similarity of two vectors, one needs to maximize their projection, that is the dot product of the two vectors. On the contrary, to maximize their orthogonality, the projection needs to be minimized. As it will be shown hereafter, by accounting the vector CSI in the scheduling process, the multiuser gains can be exploited towards further maximizing the system throughput performance.

Inspired by the multigroup multicast nature of the frame-based precoding problem, a multicast-aware user scheduling policy is developed in the present section. In the frame-based precoding methods presented in the previous sections, a precoding design over a randomly defined group of users is assumed. Since all co-scheduled users are served by the link layer mode imposed by the worst user in each group, significant performance losses from a system design perspective will be realized by this random user grouping. Acknowledging that CSI is readily available at the transmit side, since it is a requisite for the application of interference management, the optimization of the system in any required sense can be achieved by advanced scheduling methods. These methods, as shown in Fig. 4(b) and (c) are based on the exact CSI. Imperfect CSI assumptions shall be tackled in future extensions of this work.

The most intrinsic attribute of a joint scheduling and precoding design lies in the coupled nature of the two designs. Since precoding drastically affects the useful signal power at the receive side, the relation between CSI and SINR is not straightforward. The block diagram in Fig. 4(b), presents an optimal joint scheduler. This module jointly performs precoding and scheduling by feeding the output of the precoder back to the scheduler. Based on an initial user scheduling, a precoding matrix calculated by the methods of Section IV, can be applied. Then, the resulting SINR value needs to be fed back to the scheduler where a new schedule is then re-calculated. Based on this schedule, a new precoding matrix needs to be calculated and applied thus leading to a potentially different SINR distribution. Clearly, this procedure needs to be performed until all the possible combinations of users are examined. Thus, the implementation complexity of such a technique is prohibitive for

the system dimensions examined herein. A reduction of the system dimensions, on the other hand, reduces the averaging accuracy and renders the results inaccurate from a system design perspective. Therefore, the optimal user scheduling policy will not be considered for the purposes of this work.

Algorithm 3 Multicast-Aware User Scheduling Algorithm

Input: \mathbf{H}

Output: User allocation sets \mathcal{I}

begin

Step 1: $\forall l = 1, 2, \dots, N_t$ allocate semi-orthogonal users to different groups. Let $\mathcal{I} = \emptyset$ denote the index set of users allocated to groups, $\mathcal{J} = \{1, \dots, N_u\} - \{\mathcal{I}\}$ the set of unprocessed users and $g_{(1)} = \max_k \|\mathbf{h}_k\|_2$

while $|\mathcal{I}| < N_t$ **do**

forall $m \in \mathcal{J}, l = 1 \dots N_t$ **do**

$\mathbf{g}_m^\dagger = \mathbf{h}_m^\dagger \left(\mathbf{I}_{N_t} - \sum_{q=1}^l \frac{\mathbf{g}_{(q)} \mathbf{g}_{(q)}^\dagger}{\|\mathbf{g}_{(q)}\|_2^2} \right)$ calculate the orthogonal component (rejection) of each unprocessed user's channel, onto the subspace spanned by the previously selected users.

end

Select the most orthogonal user to be allocated to the l -th group: $\mathcal{G}_l = \arg \max_m \|\mathbf{g}_m\|_2$, $\mathbf{g}_{(l)} = \mathbf{g}_{\mathcal{G}_l}$ and update the user allocation sets $\mathcal{I} = \mathcal{I} \cup \{\mathcal{G}_l\}$, $\mathcal{J} = \mathcal{J} - \{\mathcal{G}_l\}$

end

Step 2: for each group select the most parallel users.

for $l = 1 \dots N_t$ **do**

while $|\mathcal{G}_l| < \rho$ **do**

forall $m \in \mathcal{J}$ **do**

$\mathbf{u}_m = \mathbf{h}_m^\dagger \frac{\mathbf{h}_j \mathbf{h}_j^\dagger}{\|\mathbf{h}_j\|_2^2}, j = [\mathcal{G}_l]_1$; calculate the projection of each users channel, onto the first user of each group. Select the user that is most parallel to the first user of each group. $\pi_l = \arg \max_m \{\|\mathbf{u}_m\|_2\}$ and update the user allocation sets $\mathcal{G}_l = \mathcal{G}_l \cup \{\pi_l\}$, $\mathcal{I} = \mathcal{I} \cup \{\mathcal{G}_l\}$, $\mathcal{J} = \mathcal{J} - \{\mathcal{G}_l\}$

end

end

end

end

As described in the previous paragraph, precoding is affected by scheduling and vice versa. To the end of providing a low complexity solution to this causality dilemma, a multicast-aware approach is illustrated in Fig. 4(c). Based on this concept, an advanced low complexity CSI based scheduling method that does not require knowledge of the resulting SINR, is developed. The key step in the proposed method lies in measuring the similarity between user channels, given the readily available CSI. The underlying intuition is that users scheduled in the same frame should have co-linear (i.e., similar) channels since they need to receive the same set of symbols (i.e., frame). On the contrary, interfering users, scheduled in adjacent synchronous frames, should be orthogonal to minimize interferences [18]. The multicast-aware user scheduling algorithm, presented in detail in Algorithm 3, is a low complexity heuristic iterative

TABLE II
LINK BUDGET PARAMETERS

Parameter	Value
Frequency Band	Ka (20 GHz)
User terminal clear sky temp, T_{cs}	235.3K
User Link Bandwidth, B_u	500 MHz
Output Back Off, OBO	5 dB
On board Power, P_{tot}	50 dBW
Roll off, α	0.20
User terminal antenna Gain, G_R	40.7 dBi
Multibeam Antenna Gain, G_{ij}	Ref: [22]

algorithm that allocates orthogonal users in different frames and simultaneously parallel users with similar channels in the same frame. In more detail, this two step algorithm operates as follows. In the first step of the process, one user per group is allocated according to the semi-orthogonality criteria originally proposed in [18]. This semi-orthogonality criterion was originally derived for zero-forcing ZF precoding, to find the users with the minimum interferences. This approach is adopted for the first step of the proposed algorithm, since the goal is to allocate non-interfering users in different groups. Next, a novel second step provides the multicast awareness of the herein proposed algorithm. In Step 2, for each of the groups sequentially, the most parallel users to the previously selected user are scheduled in the same frame. Subsequently, the similarity of the co-group channels is maximized.

VI. PERFORMANCE EVALUATION & APPLICATIONS

Based on the simulation model defined in [22], the performance of a full frequency reuse, broadband multibeam satellite that employs frame-based precoding, is compared to conventional four color reuse configurations. Since by the term user, a individual receive terminal is implied and the terms frame, beam and group are effectively equivalent in the scenario under study, the total number of users considered over the entire coverage can be found by multiplying the users per frame with the number of beams. The average user throughput given by (5) is calculated to quantify the potential gains of frame-based precoding. The rate and SINR distributions over the coverage before and after precoding are also investigated. Moreover, the sensitivity of all discussed methods to an increasing number of users per frame is presented. The simulation setup is described in Section II-A. For accurate averaging, 100 users per beam are considered uniformly distributed across the coverage area illustrated in Fig. 3. The average user throughput R_{avg} , as given via (5), is also averaged over all transmissions required to serve the initial pool of users. This consideration provides a fair comparison when user scheduling methods are considered.³ The link budget parameters considered follow the recommendations of [22] and are summarized in Table II. The minimum SINR value γ_{min} considered herein is -2.85 dB, corresponding to the minimum value supported by the normal frame operation of the most recent satcom standards [5]. Operation in even lower values is bound to increase the reported gains, since a relaxation

³Serving less users than the available for selection would drastically improve the results but not in a fair manner from a system design perspective, since this would imply that some users are denied service for an infinite time.

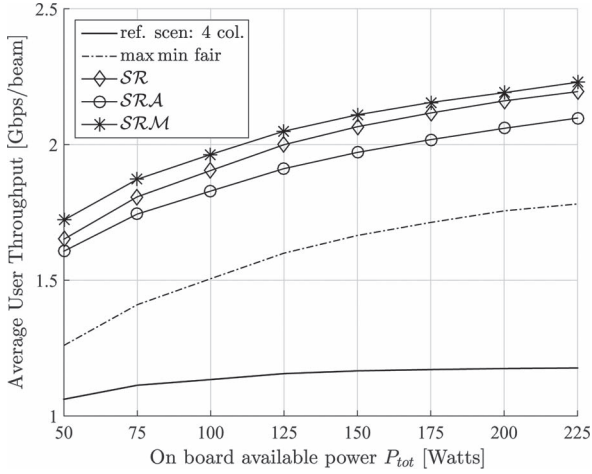


Fig. 5. Average user throughput versus on board available transmit power, for 2 users per frame.

in the added availability constraint allows for higher flexibility and thus sum-rate gains.

A. Throughput Performance

The validity of the heuristic sum-rate maximization algorithm is established by comparing the performance of the herein proposed precoders with the optimal in a max-min fair sense, solutions of [13]. The throughput versus availability tradeoff between the two formulations will also be exhibited in the following. In Fig. 5, the average user throughput of the considered multibeam satellite is plotted versus an increasing total on board available power, in [Gbps/beam]. Two users per frame are considered, i.e., $\rho = 2$. Clearly, the proposed precoding designs outperform existing approaches. The *SR* problem achieves more than 30% gains over the max min fair solutions of [13], [14]. These gains are reduced when the max SR under MRCs is considered, i.e., *SRA*. This is the price paid for guaranteeing service availability over the coverage. Finally, the maximum gains are observed when the modulation aware max SR precoding, i.e., *SRM* is employed, which also guarantees service availability. Consequently, the best performance is noted for *SRM* with more than 30% of gains over the max min fair formulation of [13] and as much as 100% gains over conventional systems in the high power region, for 2 users per frame.

For the same simulation setting, the cumulative distribution functions (CDFs) of the SINRs over the coverage area is given in Fig. 6. Clearly, conventional systems achieve higher SINRs by the means of the fractional frequency reuse. This value is around 17 dB, in line with the results of [22]. However, this does not necessarily translate to system throughput performance. To guarantee increased SINRs, the frequency allocated per user is four times reduced. On the other hand, aggressive frequency reuse reduces the average SINR values and increases its variance, as seen in Fig. 6. This, however, allows for more efficient resource utilization and consequently higher throughput, as seen before in Fig. 5. Moreover, the superiority of the max SR techniques proposed herein, over the fair solutions is also evident. Amongst these methods, the best one is *SRM* as already shown.

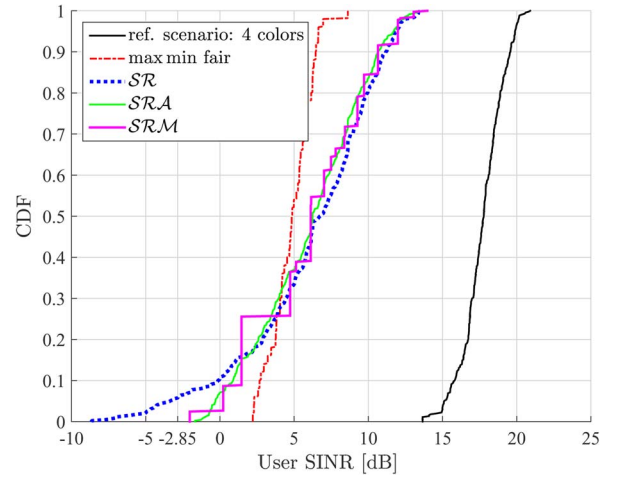


Fig. 6. CDF of user SINR over the coverage, for 2 users per frame.

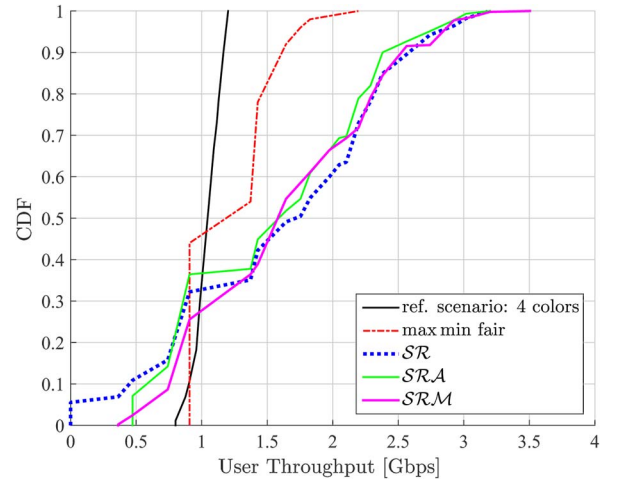


Fig. 7. Per-user throughput CDF, for 2 users per frame.

The benefits of *SRA* over *SR* are clear in Fig. 6, where the SINR CDF of all methods is presented. Clearly, *SRA* guarantees a minimum SINR of -2.85 dB but attains SINRs higher than 2 dB with less probability than *SR*. Nevertheless, *SRA* can be regarded as a middle step towards advancing to the more elaborate, *SRM* algorithm. Since *SRM* includes the same availability constraints as *SRA*, identical availability gains are noted for both methods. However, *SRM* exploits the granular nature of the spectral efficiency function towards achieving SINRs higher than *SR*. In Fig. 6, it is clear that the proposed optimization manages to adapt each user's SINR to the throughput function, since the SINR distribution follows the granular spectral efficiency function. Users have SINR values in between the DVB-S2X thresholds with very low probability. This insightful result justifies the increased gains of *SRM*, even for guaranteed availability. An additional observation from Fig. 6 is that 40% of the users operate utilizing the first four available MODCODs.

Moreover, Fig. 7 provides the rate CDFs of the conventional and the max min fair systems and exhibits the very low variance of their receive SINR. On the contrary, *SR* achieves very high rates but also drives some users to the unavailability region.

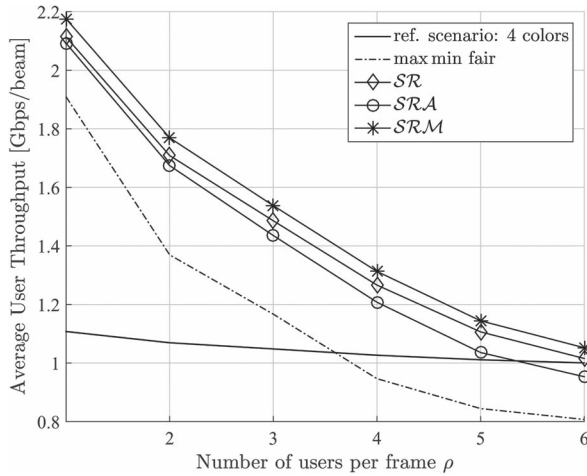


Fig. 8. Average user throughput versus the number of users per frame.

A 5% outage probability is noted for this precoding scheme. This is not the case for the *SR* and *SRM* problems, which guarantee at least 0.3 Gbps to all users.

An important issue is the performance of the developed methods with respect to an increasing number of users per frame. As presented in Fig. 8, *SRM* manages to provide more than 30% of gains for $\rho = 3$ users per frame. Both the conventional and the proposed systems suffer from an increase in the number of users per frame, since the worst user defines the MODCOD for all users. For conventional systems, this degradation is negligible when compared to the frame based precoding systems. The performance degradation when a precoding vector is matched to more than one channels is expected. As initially proven in [11], when advancing from unicasting to multicasting, the precoding problem becomes NP-hard. Added to that, when more users are grouped together, then the chances are that one of them will be compromised and thus constrain the performance of all other users. This observation further justifies the results of Fig. 8. Nevertheless, in the same figure, positive gains over the conventional systems are reported even for 6 users per frame unlike all other state of the art techniques. These results are given for a nominal on board available power of 50 Watts. It should be noted that performance in the results presented hitherto is compromised by the random user scheduling since users with very different SINRs are co-scheduled and thus constrained by the performance of the worst user.

B. Example

To the end of gaining insights on the max SR optimization, a small scale example is presented. Let us assume 2 users per frame (i.e., $\rho = 2$). The individual throughput of each user is plotted in Fig. 9 for the discussed methods. The per beam average throughput is given in the legend of the figure for each method respectively. In the conventional system, variance in the rates between the groups is noted. This results to an average user throughput equal to 1.06 Gbps/beam. By the fair optimization of [13] 1.26 Gbps/beam of are attained, while the minimum rates are balanced among the groups. More importantly, the sum rate maximizing optimization reduces the rate allocated to the

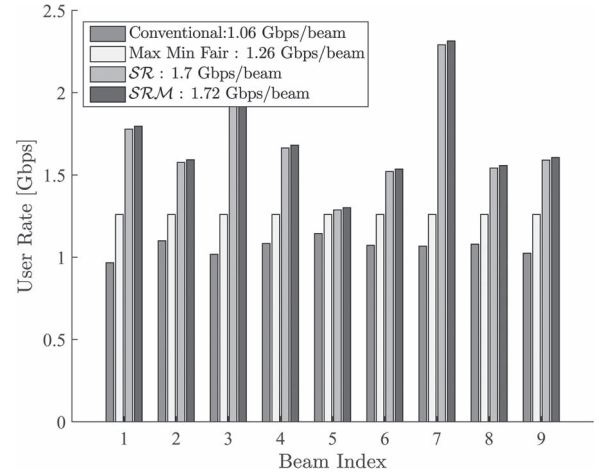


Fig. 9. Per-user achievable rate in each beam, for different designs.

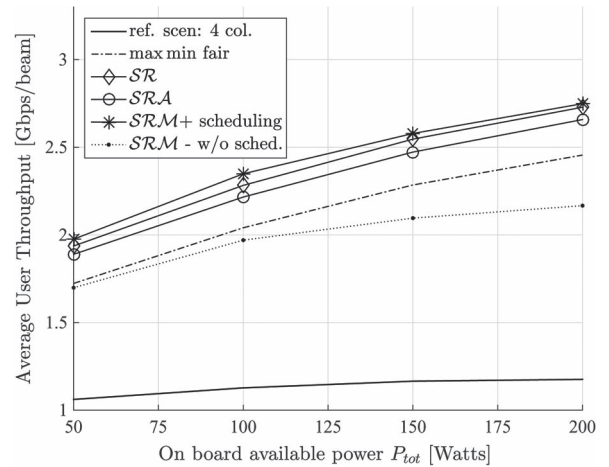


Fig. 10. Average user throughput versus on board available transmit power, for 2 users per frame, when scheduling is employed.

users in beam 5 and increases all other users. Thus, the system throughput is increased to just over 1.6 Gbps/beam. Finally, the modulation aware optimization builds upon the sum rate maximization, adapts the power allocation to the modulation constrained performances and allocates to each user equal or better rates. Consequently, it outperforms all other techniques leading to $R_{avg} = 1.72$ Gbps/beam.

C. User Scheduling

The present section presents results when the multicast-aware user scheduling algorithm is employed. In Fig. 10, the performance of the algorithm for $\rho = 2$ users per group is given versus an increasing on-board power budget. In this figure, approximately 25% of improvement the random scheduling of Section VI-A is noted. Furthermore, in Fig. 11, results for an increasing number of users per frame and for a nominal on board available power of 50 Watts, are given. The performance of *SRM* without scheduling as presented in Fig. 8, is also given for comparison. From the results of Fig. 11, it is clear that by employing user scheduling, the degradation of the system performance with respect to an increasing number of users per group is significantly improved. The same initial group of users

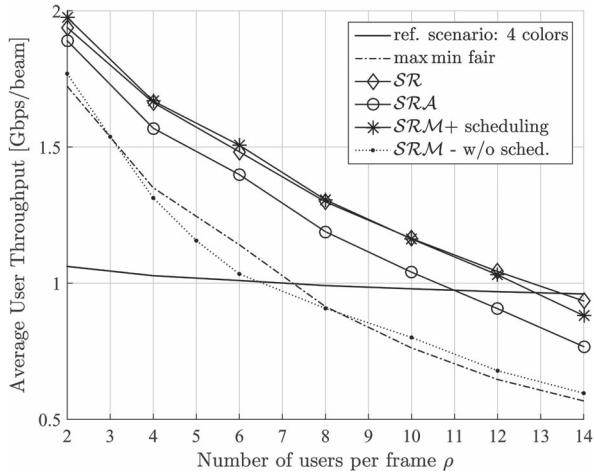


Fig. 11. Per beam throughput with respect to an increasing number of users per frame.

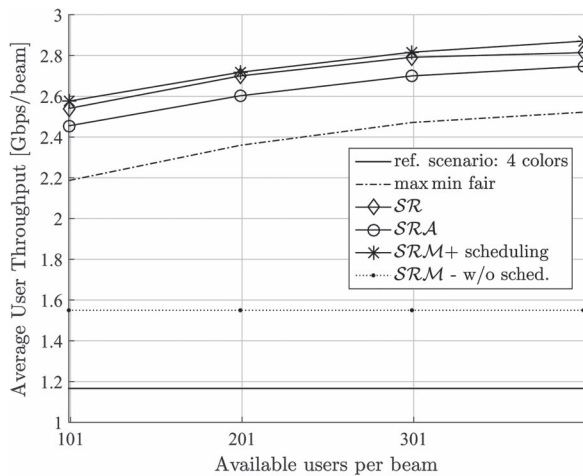


Fig. 12. Average throughput with respect to an increasing number of available users for selection users, for 3 users per frame, when scheduling is employed.

as before is employed regardless of the frame size, excluding a small rounding error cut off.⁴ The most important result is that by employing multicast-aware user scheduling methods, more than 30% of gains can be gleaned over conventional systems for as much as 7 users per frame. Also, even 13 users per frame can be accommodated in a frame with positive gains over conventional frequency reuse payload configurations. Finally, to exhibit the dependence of the performance with respect to the available for selection user pool, in Fig. 12, the average user throughput for three users per frame with respect to an increasing user pool is plotted. Almost 20% gains are noticed when doubling the user pool. Clearly, the potential of user scheduling is even higher in larger multiuser settings.

VII. CONCLUSION

In the present work, full frequency reuse configurations enabled by frame-based linear precoding are proposed for

the throughput optimization of broadband multibeam satellite systems. In this direction, sum rate optimal, frame-based precoding under per-antenna power constraints is proposed. To satisfy highly demanding in terms of availability requirements, while maintaining high gains over conventional systems, the optimization is extended to account for minimum rate and modulation constraints. Finally, to glean the satellite multiuser diversity gains, user scheduling methods adapted to the novel system design are derived. In a nutshell, the gains from frame-based precoding combined with multicast-aware user scheduling are more than 30% in terms of throughput performance, for 7 users per frame, over conventional system configurations. These gains are achieved without loss in the outage performance of the system. Also, up to 13 users per frame can be accommodated with throughput performance similar to that of the conventional systems.

Future extensions of this work include a robust frame-based precoding design to cope with CSI imperfections as well as studies to counteract the non-linearities of the satellite channel.

REFERENCES

- [1] D. Christopoulos, P.-D. Arapoglou, S. Chatzinotas, and B. Ottersten, "Linear precoding in multibeam satcoms: Practical constraints," in *Proc. 31st AIAA ICSSC*, Florence, Italy, Oct. 2013, pp. 1–9.
- [2] G. Taricco, "Linear precoding methods for multi-beam broadband satellite systems," in *Proc. 20th Eur. Wireless Conf.*, May 2014, pp. 1–6.
- [3] *Digital Video Broadcasting (DVB); Second Generation Framing Structure, Channel Coding and Modulation Systems for Broadcasting, Interactive Services, News Gathering and Other Broad-Band Satellite Applications (DVB-S2)*, European Broadcasting Union (EBU), ETSI EN 302 307-1 V1.4.1, Nov. 2014.
- [4] *Digital Video Broadcasting (DVB); Second Generation Framing Structure, Channel Coding and Modulation Systems for Broadcasting, Interactive Services, News Gathering and Other Broad-Band Satellite Applications. Part 2: DVB-S2 Extensions (DVB-S2X)*, European Broadcasting Union (EBU), ETSI EN 302 307-2 V1.1.1, Oct. 2014.
- [5] *Second Generation Framing Structure, Channel Coding and Modulation Systems for Broadcasting, Interactive Services, News Gathering and Other Broadband Satellite Applications; Part II: S2-Extensions (S2X)*, DVB Blue Book A171-2, Mar. 2014.
- [6] D. Christopoulos, S. Chatzinotas, G. Zheng, J. Grotz, and B. Ottersten, "Linear and non-linear techniques for multibeam joint processing in satellite communications," *EURASIP J. Wireless Commun. Netw.*, vol. 2012, p. 162, 2012. [Online]. Available: <http://jwcn.eurasipjournals.com/content/2012/1/162>
- [7] 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.
- [8] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Frame based precoding in satellite communications: A multicast approach," in *Proc. IEEE ASMS*, Livorno, Italy, Sep. 2014, pp. 293–299.
- [9] G. Zheng, S. Chatzinotas, and B. Ottersten, "Generic optimization of linear precoding in multibeam satellite systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 6, pp. 2308–2320, Jun. 2012.
- [10] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "User scheduling for coordinated dual satellite systems with linear precoding," in *Proc. IEEE ICC*, Budapest, Hungary, 2013, pp. 3091–3096.
- [11] N. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, Jun. 2006.
- [12] Y. C. B. Silva and A. Klein, "Linear transmit beamforming techniques for the multigroup multicast scenario," *IEEE Trans. Veh. Technol.*, vol. 58, no. 8, pp. 4353–4367, Oct. 2009.
- [13] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Weighted fair multicast multigroup beamforming under per-antenna power constraints," *IEEE Trans. Signal Process.*, vol. 62, no. 19, pp. 5132–5142, Oct. 2014.

⁴For instance, when 3 users per frame are assumed, the total number of users served is reduced to 891. This does not affect the presented results, since they are averaged over the total number of users served.

- [14] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Multicast multigroup beamforming under per-antenna power constraints," in *Proc. IEEE ICC*, Sydney, NSW, Australia, Jul. 2014, pp. 4704–4710.
- [15] M. Kaliszan, E. Pollakis, and S. Stanczak, "Multigroup multicast with application-layer coding: Beamforming for maximum weighted sum rate," in *Proc. IEEE WCNC*, 2012, pp. 2270–2275.
- [16] S. Stanczak, M. Wiczanowski, and H. Boche, *Fundamentals of Resource Allocation in Wireless Networks: Theory and Algorithms*. Manhattan, NY, USA: Springer-Verlag, Inc., 2009.
- [17] G. Dimic and N. D. Sidiropoulos, "On downlink beamforming with greedy user selection: performance analysis and a simple new algorithm," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3857–3868, Oct. 2005.
- [18] T. Yoo and A. Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, Mar. 2006.
- [19] D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Sum rate maximizing multigroup multicast beamforming under per-antenna power constraints," in *Proc. IEEE Global Commun. Conf.*, Austin, TX, USA, Dec. 2014, pp. 3354–3359.
- [20] 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.
- [21] G. Zheng, S. Chatzinotas, and B. Ottersten, "Multi-gateway cooperation in multibeam satellite systems," in *Proc. 23rd IEEE Symp. Pers. Indoor Mobile Radio Commun.*, 2012, pp. 1360–1364.
- [22] "Call of Order 2-Task 1: Fair comparison and combination of advanced interference mitigation techniques," Satellite Network of Experts (SatNEx) 3, Weßling, Germany, ESA Contract 23089/10/NL/CPL.
- [23] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal multi-user transmit beamforming: Difficult problem with a simple solution structure," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 142–148, Jul. 2014.
- [24] Z.-Q. Luo, W.-K. Ma, A.-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.
- [25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.



communications and optimization methods for multiuser MIMO communications.

Dimitrios Christopoulos (S'03–M'14) was born in Athens, Greece, in 1986. He received the Diploma degree in electrical and computer engineering from the National Technical University of Athens (NTUA), Athens, Greece, in 2010, and the Ph.D. degree in electrical engineering from the Interdisciplinary Center for Security and Trust, (SnT), University of Luxembourg, Luxembourg, in 2014. Currently he is a research associate in SnT, involved in several European research activities. His research interests include signal processing for satellite communications and optimization methods for multiuser MIMO communications.



Symeon Chatzinotas (S'06–M'09–SM'13) received the M.Eng. degree in telecommunications from Aristotle University of Thessaloniki, Greece, in 2003 and the M.Sc. and Ph.D. degrees in electronic engineering from the University of Surrey, U.K., in 2006 and 2009, respectively. He is currently a Research Scientist with the SIGCOM research group in the SnT Centre, University of Luxembourg.

Previously, he worked in numerous R&D projects for the Institute of Informatics and Telecommunications, National Center for Scientific Research, the

Institute of Telematics and Informatics, Centre of Research and Technology Hellas and the Centre of Communication Systems Research, University of Surrey. He has authored more than 150 technical papers in refereed international journals, conferences, and scientific books. He has coedited a book titled *Cooperative and Cognitive Satellite Systems*. His research interests are on multiuser information theory, cooperative/cognitive communications, and wireless networks optimization. He is co-organizing the First Workshop on "Cognitive Radios and Networks for Spectrum Coexistence of Satellite and Terrestrial Systems" in conjunction with the IEEE ICC 2015, London, U.K. He is a corecipient of the 2014 Distinguished Contributions to Satellite Communications Award from the IEEE Communications Society and coauthor of the 2015 CROWNCOM Best Paper Award.



Björn Ottersten (S'87–M'89–SM'99–F'04) was born in Stockholm, Sweden, in 1961. He received the M.S. degree in electrical engineering and applied physics from Linköping University, Linköping, Sweden, in 1986 and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1989.

Dr. Ottersten has held research positions at the Department of Electrical Engineering, Linköping University, the Information Systems Laboratory, Stanford University, the Katholieke Universiteit

Leuven, Leuven, and the University of Luxembourg. During 1996 to 1997, he was Director of Research at ArrayComm Inc, a start-up in San Jose, CA, USA, based on Ottersten's patented technology. He has co-authored journal papers that received the IEEE Signal Processing Society Best Paper Award in 1993, 2001, 2006, and 2013 and 3 IEEE conference papers receiving Best Paper Awards. In 1991, he was appointed Professor of Signal Processing at the Royal Institute of Technology (KTH), Stockholm. From 1992 to 2004, he was Head of the department for Signals, Sensors, and Systems at KTH and from 2004 to 2008, he was Dean of the School of Electrical Engineering at KTH. Currently, he is Director for the Interdisciplinary Centre for Security, Reliability and Trust at the University of Luxembourg. As Digital Champion of Luxembourg, he acts as an adviser to European Commissioner Neelie Kroes.

Dr. Ottersten has served as Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and on the editorial board of *IEEE Signal Processing Magazine*. He is currently editor in chief of *EURASIP Signal Processing Journal* and a member of the editorial boards of *EURASIP Journal of Applied Signal Processing* and *Foundations and Trends in Signal Processing*. He is a Fellow of the IEEE and EURASIP and a member of the IEEE Signal Processing Society Board of Governors. In 2011, he received the IEEE Signal Processing Society Technical Achievement Award. He is a first recipient of the European Research Council advanced research grant. His research interests include security and trust, reliable wireless communications, and statistical signal processing.