Joint Optimization of Power and Data Transfer in Multiuser MIMO Systems

Javier Rubio, Antonio Pascual-Iserte, *Senior Member, IEEE*, Daniel P. Palomar, *Fellow, IEEE*, and Andrea Goldsmith, *Fellow, IEEE*

***Abstract*—We present an approach to solve the nonconvex optimization problem that arises when designing the transmit covariance matrices in multiuser multiple-input multiple-output (MIMO) broadcast networks implementing simultaneous wireless information and power transfer (SWIPT). The MIMO SWIPT problem is formulated as a general multiobjective optimization problem, in which data rates and harvested powers are optimized simultaneously. Two different approaches are applied to reformulate the (nonconvex) multiobjective problem. In the first approach, the transmitter can control the specific amount of power to be harvested by power transfer whereas in the second approach the transmitter can only control the proportion of power to be harvested among the different harvesting users. We solve the resulting formulations using the majorization–minimization (MM) approach. The solution obtained from the MM approach is compared to the classical block-diagonalization (BD) strategy, typically used to solve the nonconvex multiuser MIMO network by forcing no interference among users. Simulation results show that the proposed approach improves over the BD approach both the system sum rate and the power harvested by users. Additionally, the computational times needed for convergence of the proposed methods are much lower than the ones required for classical gradient-based approaches.**

***Index Terms*—Energy harvesting, power transfer, SWIPT, majorization minimization, MIMO, nonconvex optimization.**

I. Introduction

SIMULTANEOUS wireless information and power transfer (SWIPT) is a transmission technique in which a transmitter actively feeds a receiver (or a set of receivers) power that is sent through radio frequency (RF) signals and, simultaneously, communicates information to the same or different set of receivers [1]. Battery-constrained devices are able to recharge their batteries by collecting the transmitted power and, thus, enhance their operation time [2]. Currently, there are different energy harvesting techniques that can be used to power devices, such as wind or solar, but SWIPT technology represents an appealing solution as the transmitter is able to control explicitly the amount of energy that the device will receive and, hence, keep them alive. Historically, due to the high attenuation of signals over distance, SWIPT techniques were only introduced in low-power devices, such as RFID tags [3]. However, new advances in hardware technologies have enabled power to be transferred and harvested much more efficiently [1], [3].

The first paper in the literature that covered the concept of SWIPT is the one by Varshney [4]. He showed that there exists a nontrivial trade-off in maximizing the data rate with power transmission constraints. Zhang and Ho [5] developed a SWIPT technique for multiple-input multiple-output (MIMO) scenario, composed of one transmitter capable of transmitting information and power simultaneously to one receiver. Then, Rubio and Pascual-Iserte [6], extended the work in [5] by considering that multiple users were present in the MIMO system. But since the multi-stream transmit covariance optimization that arises in SWIPT MIMO systems is a very difficult nonconvex optimization problem, they considered a block-diagonalization (BD) strategy [7] in which interference is pre-canceled at the transmitter. The BD technique allows for a simple solution but wastes some degrees of freedom and, thus, the performance obtained may be lower than the one obtained by solving the nonconvex problem. Works [8] and [9] considered a MIMO network consisting of multiple transmitter-receiver pairs with co-channel interference. The study in [8] focused on the case with two transmitter-receiver pairs whereas in [9], the authors generalized [8] by considering that k transmitter-receivers pairs were present. The work in [10] considered a MIMO system with single-stream transmission, with the objective of minimizing the overall power consumption with per-user signal to interference and noise ratio (SINR) constraints and harvesting constraints. The design of multiuser broadcast networks under the framework of multiple-input single-output (MISO) beamformimg optimization has also been addressed in works such as [11] and [12].

There exist two approaches in the literature that deal with the nonconvex optimization of the transmit covariance matrices in multiuser multi-stream MIMO networks. The first is based on the duality principle [13]. In [14], Gui *et al.* applied that principle to obtain the beamforming optimization solution for the multiuser MIMO SWIPT broadcast channel. However,

Manuscript received April 2, 2016; revised July 31, 2016; accepted August 23, 2016. Date of publication October 3, 2016; date of current version November 4, 2016. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Yue Rong. This work was supported in part by the European Commission in the framework of the FP7 Network of Excellence in Wireless COMmunications NEWCOM# under Grant 318306, in part by the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competitividad) under Project TEC2011-29006-C03-02 (GRE3N-LINK-MAC), Project TEC2013-41315-R (DISNET), and FPI Grant BES-2012-052850, in part by the Catalan Government (AGAUR) under Grant 2014 SGR 60, in part by the Hong Kong Government under Research Grant Hong Kong RGC 16207814 and 16206315, and in part by the NSF Center for Science of Information (CSoI): NSF-CCF-0939370.

J. Rubio and A. Pascual-Iserte are with the Universitat Politècnica de Catalunya, Barcelona 08034, Spain (e-mail: javier.rubio.lopez@upc.edu; antonio.pascual@upc.edu).

D. P. Palomar is with the Hong Kong University of Science and Technology, Clear Water Bay Hong Kong (e-mail: palomar@ust.hk).

A. Goldsmith is with Stanford University, Stanford, CA 94305 USA (e-mail: andrea@wsl.stanford.edu).

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/TSP.2016.2614794

that work considered an overall (sum) harvesting constraint instead of individual per-user harvesting constraints. The second approach is based on the minimization of the mean square error (MSE) [15]. However, this technique cannot be applied to the SWIPT framework due to fact that the resulting problem remains nonconvex.

The main difference of our work with respect to the previous works described above is that we assume a broadcast multiuser multi-stream (non BD-based) MIMO SWIPT network, in which (per-user) harvested power and information transfer must be optimized simultaneously. We model our transmitter design as a multi-objective problem in which the scenarios studied in [5] and [6] are shown to be particular solutions of the proposed framework. Additionally, we assume that interference is not precanceled (i.e., the BD approach is not applied) and, thus, both larger information transfer and harvested power can be achieved simultaneously. The resulting problem is nonconvex and very difficult to solve. In order to obtain local solutions, we derive different methods based on majorization-minimization (MM) techniques. By means of this strategy, we are able to reformulate our original nonconvex problem into a series of convex subproblems that are easily solved (i.e., through algorithms that have a very low computational complexity) and whose solutions converge to a locally optimal solution of the original nonconvex problem.

The techniques based on MM that we propose in this journal paper are also compared in the simulations section with other previous algorithms used as benchmarks and listed in Section IV.C. Some of these algorithms used as benchmarks were developed by the same authors and presented in the conference paper [16].

The remainder of this paper is organized as follows. In Section II, we introduce a summary of the mathematical techniques employed in this paper. In Section III we present the system and signal models and the problem formulation. In Section IV we derive the mathematical modeling required to reformulate the original nonconvex problem into convex subproblems that are solved using the MM approach. In Section V, we evaluate the performance of the proposed methods and, finally, in Section VI, we draw some conclusions.

Notation: We adopt the notation of using boldface lower case for vectors $\mathbf{x}$ and upper case for matrices $\mathbf{X}$. The transpose, conjugate transpose (hermitian), and inverse operators are denoted by the superscripts $(\cdot)^T$, $(\cdot)^H$, and $(\cdot)^{-1}$, respectively. $\mathrm{Tr}(\cdot)$ and $\det(\cdot)$ denote the trace and the determinant of a matrix, respectively. $\mathrm{vec}(\mathbf{X})$ is a column vector resulting from stacking all columns of $\mathbf{X}$. We use $\mathbf{X}$ to denote the $N-$tuple $\mathbf{X} \triangleq (\mathbf{X}_i)_{i=1}^N = (\mathbf{X}_1, \ldots, \mathbf{X}_N)$ and $||\cdot||_F$ to denote the matrix Frobenius norm.

II. Mathematical Preliminaries

A. Multi-Objective Optimization

Multi-objective optimization (also known as multi-criteria optimization or vector optimization) is a type of optimization that involves multiple objective functions that are optimized simultaneously [17]. For a nontrivial multi-objective problem, in general, there does not exist a single solution that simultaneously optimizes each objective. In that case, the objective functions are said to be conflicting, and there exists a (possibly infinite) number of Pareto optimal solutions. A solution is called Pareto optimal if none of the objective functions can be improved in value without degrading some of the other objective values.

1) Definitions:

Definition 1 ([17]): A multi-objective problem can be formally expressed as

$$\begin{aligned} \underset{\boldsymbol{x}}{\text{maximize}} \quad & \boldsymbol{f}(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_K(\boldsymbol{x})) \\ \text{subject to} \quad & \boldsymbol{x} \in \mathcal{X}, \end{aligned} \tag{1}$$

where $f_k : \mathbb{C}^N \to \mathbb{R}$ for $k = 1, \ldots, K$ and $\mathcal{X}$ is the feasible set that represents the constraints. Let $\mathcal{Y}$ be the set of all attainable points for all feasible solutions, i.e., $\mathcal{Y} = \mathbf{f}(\mathcal{X})$.

2) Efficient Solutions:

Definition 2 ([17], Definition 2.1): A point $\mathbf{x} \in \mathcal{X}$ is called Pareto optimal if there is no other $\mathbf{x}' \in \mathcal{X}$ such that $\mathbf{f}(\mathbf{x}') \succeq \mathbf{f}(\mathbf{x})$, where $\succeq$ refers to the component-wise inequality, i.e., $f_i(\mathbf{x}') \geq f_i(\mathbf{x})$, $i = 1, \ldots, K$.

Sometimes, ensuring Pareto optimality for some problems is difficult. Due to this, the condition of optimality can be relaxed as follows.

Definition 3 ([17], Definition 2.24): A point $\mathbf{x} \in \mathcal{X}$ is called weakly Pareto optimal (or weakly efficient) if there is no other $\mathbf{x}' \in \mathcal{X}$ such that $\mathbf{f}(\mathbf{x}') \succ \mathbf{f}(\mathbf{x})$, where $\succ$ refers to the strict component-wise inequality, i.e., $f_i(\mathbf{x}') > f_i(\mathbf{x})$, $i = 1, \ldots, K$. All Pareto optimal solutions are also weakly Pareto optimal.

3) Finding Pareto Optimal Points: There are several methods for finding the Pareto points of a multi-objective problem. In the sequel, we present three different (scalarization) techniques.

a) Weighted sum method: The simplest scalarization technique is the weighted sum method which collapses the vector-objective into a single-objective component sum:

$$\underset{\mathbf{x} \in \mathcal{X}}{\text{maximize}} \quad \sum_{k=1}^{K} \beta_k f_k(\mathbf{x}), \tag{2}$$

where β_k are real non-negative weights. The following results present the relation between the optimal solutions of (2) and the Pareto optimal points of the original problem (1).

Proposition 1 ([17], Proposition 3.9): Suppose that $\mathbf{x}^\star$ is an optimal solution of (2). Then, $\mathbf{x}^\star$ is weakly efficient.

Proposition 2 ([17], Proposition 3.10): Let $\mathcal{X}$ be a convex set, and let f_k be concave functions, $k = 1, \ldots, K$. If $\mathbf{x}^\star$ is weakly efficient, there are some $\beta_k \geq 0$ such that $\mathbf{x}^\star$ is an optimal solution of (2).

As a result, convexity is apparently required for finding all weakly Pareto optimal points with the weighted sum method, which means that if the original problem is not convex, all the Pareto optimal points may not be found by using the weighted sum method. However, there are other weighted sum techniques in the literature (see, for example, the adaptive weighted sum method [18]) that are able to find all Pareto optimal points for nonconvex problems at the expense of a higher computational complexity.

b) Epsilon-constraint method: In this method, only one of the original objectives is maximized while the others are transformed into constraints:

$$\begin{aligned} &\underset{\mathbf{x}\in\mathcal{X}}{\text{maximize}} \quad f_j(\mathbf{x}) \\ &\text{subject to} \quad f_k(\mathbf{x}) \geq \epsilon_k, \quad k=1,\ldots,K,\ k\neq j. \end{aligned} \tag{3}$$

Let us introduce the following results.

Proposition 3 ([17], Proposition 4.3): Let $\mathbf{x}^\star$ be an optimal solution of (3) for some j. Then $\mathbf{x}^\star$ is weakly Pareto optimal.

Proposition 4 ([17], Proposition 4.5): A feasible solution $\mathbf{x}^\star \in \mathcal{X}$ is Pareto optimal if, and only if, there exists a set of $\hat{\epsilon}_k, k=1,\ldots,K$ such that $\mathbf{x}^\star$ is an optimal solution of (3) for all $j=1,\ldots,K$.

Contrary to the weighted sum method, convexity is not needed in the previous two propositions (but convexity is still typically required to solve problems like (3)).

c) Hybrid method: This method combines the previous two methods, i.e., the weighted sum method and the epsilon-constraint method. In this case, the scalarized problem to be solved has a weighted sum objective and constraints on all (or some) objectives as follows:

$$\begin{aligned} &\underset{\mathbf{x}\in\mathcal{X}}{\text{maximize}} \quad \sum_{k\in\mathcal{K}_1} \beta_k f_k(\mathbf{x}) \\ &\text{subject to} \quad f_k(\mathbf{x}) \geq \epsilon_k, \quad k \in \mathcal{K}_2, \end{aligned} \tag{4}$$

where $|\mathcal{K}_1| \leq K$,f $|\mathcal{K}_2| \leq K$, for $|\mathcal{A}|$ the cardinality of set $\mathcal{A}$, and β_k are real non-negative weights.

B. Majorization-Minimization Method

The MM is an approach to solve optimization problems that are too difficult to solve in their original formulation. The principle behind the MM method is to transform a difficult problem into a sequence of simple problems. Interested readers may refer to [19] and the references therein for more details.

The method works as follows. Suppose that we want to maximize $f_0(\mathbf{x})$ over $\mathcal{X}$. In the MM approach, instead of maximizing the cost function $f_0(\mathbf{x})$ directly, the algorithm optimizes a sequence of approximate objective functions that minorize $f_0(\mathbf{x})$, producing a sequence $\{\mathbf{x}^{(k)}\}$ according to the following update rule:

$$\mathbf{x}^{(k+1)} = \underset{\mathbf{x}\in\mathcal{X}}{\arg\max}\ \hat{f}_0(\mathbf{x}, \mathbf{x}^{(k)}), \tag{5}$$

where $\mathbf{x}^{(k)}$ is the point generated by the algorithm at iteration k and $\hat{f}_0(\mathbf{x}, \mathbf{x}^{(k)})$, known as a surrogate function, is the minorization function of $f_0(\mathbf{x})$ at $\mathbf{x}^{(k)}$, i.e., it has to be a global lower bound tight at $\mathbf{x}^{(k)}$. Problem (5) will be referred to as the surrogate problem of the overall maximization problem (i.e., maximize $f_0(\mathbf{x})$ over $\mathcal{X}$). In addition, the surrogate function must also be continuous in $\mathbf{x}$ and $\mathbf{x}^{(k)}$. The last condition that the surrogate function must fulfill is that its directional derivatives[1] and of the original objective function $f_0(\mathbf{x})$ must be equal at the point $\mathbf{x}^{(k)}$. All in all, the four conditions for the surrogate function are as follows:

$$(\text{A1}): \quad \hat{f}_0(\mathbf{x}^{(k)}, \mathbf{x}^{(k)}) = f_0(\mathbf{x}^{(k)}), \quad \forall \mathbf{x}^{(k)} \in \mathcal{X}, \tag{6}$$

$$(\text{A2}): \quad \hat{f}_0(\mathbf{x}, \mathbf{x}^{(k)}) \leq f_0(\mathbf{x}), \quad \forall \mathbf{x}, \mathbf{x}^{(k)} \in \mathcal{X}, \tag{7}$$

$$(\text{A3}): \quad \hat{f}_0'(\mathbf{x}, \mathbf{x}^{(k)}; \mathbf{d})|_{\mathbf{x}=\mathbf{x}^{(k)}} = f_0'(\mathbf{x}^{(k)}; \mathbf{d}), \quad \forall \mathbf{d} \text{ with } \mathbf{x}^{(k)} + \mathbf{d} \in \mathcal{X}, \tag{8}$$

$$(\text{A4}): \quad \hat{f}_0(\mathbf{x}, \mathbf{x}^{(k)}) \text{ is continuous in } \mathbf{x} \text{ and } \mathbf{x}^{(k)}. \tag{9}$$

Under assumptions (A1)–(A4), every limit point of the sequence $\{\mathbf{x}^{(k)}\}$ is a locally optimal point of the original problem (globally optimal if the problem is convex) (see [19] for details).

III. Problem Formulation

Let us consider a wireless broadcast multiuser system consisting of one base station (BS) transmitter equipped with n_T antennas and a set of K receivers, denoted as $\mathcal{U}_T = \{1, 2, \ldots, K\}$, where the k-th receiver is equipped with n_{R_k} antennas [20]. We assume that a given user is not able to decode information and to harvest energy simultaneously, and that a user being served with information by the BS uses all the energy to decode the signal. Thus, the set of users is partitioned into two disjoint subsets. One that contains the information users, denoted as $\mathcal{U}_I \subseteq \mathcal{U}_T$ with $|\mathcal{U}_I| = N$, and the other subset that contains harvesting users, denoted as $\mathcal{U}_E \subseteq \mathcal{U}_T$ with $|\mathcal{U}_E| = M$. Therefore, $\mathcal{U}_I \cap \mathcal{U}_E = \emptyset$ and $|\mathcal{U}_I| + |\mathcal{U}_E| = N + M = K$.[2] Without loss of generality (w.l.o.g.), let us index users as $\mathcal{U}_I = \{1, \ldots, N\}$ and $\mathcal{U}_E = \{N+1, \ldots, N+M\}$.

The equivalent baseband channel from the BS to the k-th receiver is denoted by $\mathbf{H}_k \in \mathbb{C}^{n_{R_k} \times n_T}$. It is also assumed that the set of matrices $\{\mathbf{H}_k\}$ is known to the BS and to the corresponding receivers (the case of imperfect CSI is outside the scope of the paper).

As far as the signal model is concerned, the received signal for the i-th information receiver can be modeled as

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{B}_i \mathbf{x}_i + \mathbf{H}_i \sum_{\substack{k \in \mathcal{U}_I \\ k \neq i}} \mathbf{B}_k \mathbf{x}_k + \mathbf{n}_i, \quad \forall i \in \mathcal{U}_I. \tag{10}$$

In the previous notation, $\mathbf{B}_i \mathbf{x}_i$ represents the transmitted signal for user $i \in \mathcal{U}_I$, where $\mathbf{B}_i \in \mathbb{C}^{n_T \times n_{S_i}}$ is the precoder matrix and $\mathbf{x}_i \in \mathbb{C}^{n_{S_i} \times 1}$ represents the information symbol vector. It is also assumed that the signals transmitted to different users are independent and zero mean. n_{S_i} denotes the number of streams assigned to user $i \in \mathcal{U}_I$ and we assume that $n_{S_i} = \min\{n_{R_i}, n_T\}\ \forall i \in \mathcal{U}_I$. The transmit covariance matrix is $\mathbf{S}_i = \mathbf{B}_i \mathbf{B}_i^H$ if we assume w.l.o.g. that $\mathbb{E}\left[\mathbf{x}_i \mathbf{x}_i^H\right] = \mathbf{I}_{n_{S_i}}$. $\mathbf{n}_i \in \mathbb{C}^{n_{R_i} \times 1}$ denotes the receiver noise vector, which is

[1] Let $f : \mathbb{C}^N \to \mathbb{R}$. Then, the directional derivative of $f(\mathbf{x})$ in the direction of vector $\mathbf{d}$ is given by $f'(\mathbf{x}; \mathbf{d}) \triangleq \lim_{\lambda \to 0} \frac{f(\mathbf{x}+\lambda\mathbf{d}) - f(\mathbf{x})}{\lambda}$.

[2] In this paper, we assume for simplicity in the formulation that a user belongs to either the harvesting set or the information set and that both sets are known and fixed. This assumption could be generalized by considering that some users are not selected in either set as well as by defining which particular users are scheduled in each particular set (i.e., user grouping strategies). However, this falls out of the scope of this paper.

considered Gaussian with $\mathbb{E}\left[\mathbf{n}_i\mathbf{n}_i^H\right] = \mathbf{I}_{n_{R_i}}$.[3] Note that the middle term of (10) is an interference term. The covariance matrix of the interference plus noise is written as

$$\mathbf{\Omega}_i(\mathbf{S}_{-i}) = \mathbf{H}_i\mathbf{S}_{-i}\mathbf{H}_i^H + \mathbf{I}, \quad \forall i \in \mathcal{U}_I, \tag{11}$$

where $\mathbf{S}_{-i} = \sum_{\substack{k\in\mathcal{U}_I\\k\neq i}} \mathbf{S}_k$. Let $\tilde{\mathbf{x}} = \mathbf{B}\mathbf{x}$ denote the signal vector transmitted by the BS, where the joint precoding matrix is defined as $\mathbf{B} = [\mathbf{B}_1 \quad \ldots \quad \mathbf{B}_N] \in \mathbb{C}^{n_T\times n_S}$, for $n_S = \sum_{i\in\mathcal{U}_I} n_{S_i}$ the total number of streams of all information users, and the data vector $\mathbf{x} = \left[\mathbf{x}_1^T \quad \ldots \quad \mathbf{x}_N^T\right]^T \in \mathbb{C}^{n_S\times 1}$, that must satisfy the power constraint formulated as $\mathbb{E}[\|\tilde{\mathbf{x}}\|^2] = \sum_{i\in\mathcal{U}_I}\mathrm{Tr}(\mathbf{S}_i) \leq P_T$, where P_T represents the total available transmission power at the BS.

The total RF-band power harvested by the j-th user from all receiving antennas, denoted by $\bar{Q}_j$, is proportional to that of the equivalent baseband signal[4], i.e., $\forall j \in \mathcal{U}_E$, we have:

$$\bar{Q}_j = \zeta_j\mathbb{E}\left[\left\|\mathbf{H}_j\sum_{i\in\mathcal{U}_I}\mathbf{B}_i\mathbf{x}_i\right\|^2\right] = \zeta_j\sum_{i\in\mathcal{U}_I}\mathbb{E}[\|\mathbf{H}_j\mathbf{B}_i\mathbf{x}_i\|^2], \tag{12}$$

where ζ_j is a constant that accounts for the loss for converting the harvested RF power to electrical power. Notice that, for simplicity, in (12) we have omitted the harvested power due to the noise term since it can be assumed negligible.

The transmitter design that we propose in this paper is modeled as a nonconvex multi-objective optimization problem. The goal is to maximize, simultaneously, the individual data rates and the harvested powers of the information and harvesting users, respectively. Given this and the previous system model, the optimization problem is written as

$$\begin{aligned}\underset{\{\mathbf{S}_i\}}{\text{maximize}} \quad & \left((R_n(\mathbf{S}))_{n\in\mathcal{U}_I}, (E_m(\mathbf{S}))_{m\in\mathcal{U}_E}\right) \\ \text{subject to} \quad & C1: \sum_{i\in\mathcal{U}_I}\mathrm{Tr}(\mathbf{S}_i) \leq P_T \\ & C2: \mathbf{S}_i \succeq 0, \quad \forall i \in \mathcal{U}_I,\end{aligned} \tag{13}$$

where $\mathbf{S} \triangleq (\mathbf{S}_i)_{\forall i\in\mathcal{U}_I}$, the data rate expression is given by

$$\begin{aligned}R_n(\mathbf{S}) &= \log\det\left(\mathbf{I} + \mathbf{H}_n\mathbf{S}_n\mathbf{H}_n^H\mathbf{\Omega}_n^{-1}(\mathbf{S}_{-n})\right) && (14)\\ &= \log\det\left(\mathbf{\Omega}_n(\mathbf{S}_{-n}) + \mathbf{H}_n\mathbf{S}_n\mathbf{H}_n^H\right) \\ &\quad - \log\det\left(\mathbf{\Omega}_n(\mathbf{S}_{-n})\right) && (15)\\ &= \underbrace{\log\det\left(\mathbf{I} + \mathbf{H}_n\bar{\mathbf{S}}\mathbf{H}_n^H\right)}_{\triangleq s_n(\mathbf{S})} - \underbrace{\log\det\left(\mathbf{\Omega}_n(\mathbf{S}_{-n})\right)}_{\triangleq g_n(\mathbf{\Omega}_n(\mathbf{S}_{-n}))}, && (16)\end{aligned}$$

with $\bar{\mathbf{S}} = \sum_{k\in\mathcal{U}_I}\mathbf{S}_k$, and the harvested power is given by

$$E_m(\mathbf{S}) = \sum_{i\in\mathcal{U}_I}\mathrm{Tr}(\mathbf{H}_m\mathbf{S}_i\mathbf{H}_m^H). \tag{17}$$

[3]We assume that noise power $\sigma^2 = 1$ w.l.o.g., otherwise we could simply apply a scale factor at the receiver and re-scale the channels accordingly.

[4]In this paper we assume that the harvested power is proportional to that of the received baseband signal. However, in work [21] authors consider a nonlinear model for the harvested power that better captures the practical energy harvesting circuits. The application of nonlinear models is out of the scope of this paper and is left as a future work.

The previous problem in (13) is not convex due the objective functions (in fact, due to $\mathbf{\Omega}_i(\mathbf{S}_{-i})$) and is difficult to solve. In order to find Pareto optimal points, we can reformulate it by using any of the techniques presented in Section II-A. In the following, we propose two approaches based on the weighted sum method and on the hybrid method. For convenience, we start with the hybrid method as it is the one that has received the most attention in the literature [5], [22]. However in that literature, the interference in (11) is assumed to be removed by the transmission strategy. This assumption makes the problem convex and hence easier to solve.

A. Hybrid-Based Formulation to Solve (13)

In the hybrid approach, some of the objective functions are collapsed into a single objective by means of scalarization and some of the objective functions are added as constraints. In particular, the data rates are left in the objective whereas the harvesting constraints are included as individual harvesting constraints. With this particular formulation, we are able to guarantee a minimum value for the power to be harvested by the harvesting users. Thus, problem (13) is formulated as

$$\begin{aligned}\max_{\{\mathbf{S}_i\}} \quad & \sum_{i\in\mathcal{U}_I}\omega_i\log\det\left(\mathbf{I} + \mathbf{H}_i\bar{\mathbf{S}}\mathbf{H}_i^H\right) - \omega_i\log\det\left(\mathbf{\Omega}_i(\mathbf{S}_{-i})\right) \\ \text{s.t.} \quad & C1: \sum_{i\in\mathcal{U}_I}\mathrm{Tr}(\mathbf{H}_j\mathbf{S}_i\mathbf{H}_j^H) \geq Q_j, \quad \forall j \in \mathcal{U}_E \\ & C2: \sum_{i\in\mathcal{U}_I}\mathrm{Tr}(\mathbf{S}_i) \leq P_T \\ & C3: \mathbf{S}_i \succeq 0, \quad \forall i \in \mathcal{U}_I,\end{aligned} \tag{18}$$

where $Q_j = \frac{\bar{Q}_j^{\min}}{\zeta_j}$, being $\{\bar{Q}_j^{\min}\}$ the set of minimum power harvesting constraints, and ω_i are some real non-negative weights. For simplicity in the notation, let us define the feasible set $\mathcal{S}_1$ as

$$\mathcal{S}_1 \triangleq \left\{\mathbf{S}: \sum_{i\in\mathcal{U}_I}\mathrm{Tr}(\mathbf{H}_j\mathbf{S}_i\mathbf{H}_j^H) \geq Q_j, \forall j \in \mathcal{U}_E, \right. \left. \sum_{i\in\mathcal{U}_I}\mathrm{Tr}(\mathbf{S}_i) \leq P_T, \mathbf{S}_i \succeq 0, \forall i \in \mathcal{U}_I\right\}. \tag{19}$$

For a set of fixed harvesting constraints, the convex hull of the rate region can be obtained by varying the values of ω_i. In addition, we can use the values of the weights to assign priorities to some users if user scheduling is to be implemented, following, for example, the proportional fair criterion [23], [24]. Notice that constraint $C1$ is associated with the minimum power to be harvested for a given user. Note also the similarities of problem (18) with the single user case presented in [5] and its extension to the multiuser case presented in [6]. As commented before, the novelty is that we do not force the transmitter to cancel the interference generated among the information users (as opposed to BD approaches [7]) and, thus, we allow the system to have more degrees of freedom to improve the system throughput and the harvested power simultaneously. Later in Section IV-A, we will present a method based on MM to solve the nonconvex problem in (18).

B. Weighted Sum-Based Formulation to Solve (13)

In situations where the exact amount of power to be harvested by harvesting users is not needed, we can also obtain Pareto optimal points by means of the simpler weighted-sum method. In this case, we can assign priorities so that some users tend to harvest more power than others, although the exact amounts cannot be controlled. As we will see later, the overall problem based on this new formulation is much easier to solve. The transmitter design is obtained through the following nonconvex optimization problem:

$$\begin{aligned}\max_{\{\mathbf{S}_i\}} \quad & \sum_{i\in\mathcal{U}_I} \omega_i \log\det\left(\mathbf{I}+\mathbf{H}_i\bar{\mathbf{S}}\mathbf{H}_i^H\right) - \omega_i \log\det\left(\boldsymbol{\Omega}_i(\mathbf{S}_{-i})\right) \\ & + \sum_{j\in\mathcal{U}_E}\sum_{i\in\mathcal{U}_I} \alpha_j \operatorname{Tr}(\mathbf{H}_j\mathbf{S}_i\mathbf{H}_j^H) \qquad (20)\\ \text{s.t.} \quad & C1: \sum_{i\in\mathcal{U}_I}\operatorname{Tr}(\mathbf{S}_i) \le P_T \\ & C2: \mathbf{S}_i \succeq 0, \quad \forall i\in\mathcal{U}_I,\end{aligned}$$

where α_j are some real non-negative weights. For simplicity in the notation, let us define the feasible set $\mathcal{S}_2$ as

$$\mathcal{S}_2 \triangleq \left\{\mathbf{S} : \sum_{i\in\mathcal{U}_I}\operatorname{Tr}(\mathbf{S}_i)\le P_T,\, \mathbf{S}_i\succeq 0, \forall i\in\mathcal{U}_I\right\}. \qquad (21)$$

As we will show later in Section IV-B, the algorithm to solve (20) is easier than the algorithm to solve (18). Hence, there is a trade-off in terms of speed of convergence of the algorithms and in terms of the harvested power control since, as we introduced before, in (18) the transmitter can fully control the amount of power to be harvested by the users whereas in (20) the transmitter can only control the proportion of the power to be harvested among the users.

IV. MM-Based Techniques to Solve Problem (13)

In this section, we present a method based on the MM philosophy to solve problems (18) and (20). Since the original problems (18) and (20) are nonconvex, we reformulate them and make them convex before applying the MM method. This reformulation will follow two steps. In the first step, problems (18) and (20) will be convexified by using a linear approximation of the nonconvex terms. This is the approach taken in papers such as [25], [26], and [27]. Instead of solving the reformulated (convex) problem, in the second step, we design a quadratic approximation of the remaining convex terms in order to find a surrogate problem easier to solve. Finally, we apply the MM method to the quadratic reformulation.

As benchmarks for comparison, we will consider the case of just convexifying the nonconvex terms, which is an approach taken in the previous literature, and also consider a gradient method applied directly to the nonconvex problems (18) and (20).

Although the mathematical developments of the proposed MM approaches are more tedious than the approaches usually taken in the literature, the resulting algorithms are faster.

A. Approach to Solve the Hybrid-Based Formulation in (18)

As we introduced before, we need to reformulate the original nonconvex problem (18) and make it convex. This will be done in two steps. Motivated by the work in [26], in this first step we derive a linear approximation for the nonconcave (right-hand side) part of the objective function of (18), i.e., $f_0(\mathbf{S}) = \sum_{i\in\mathcal{U}_I}\omega_i s_i(\mathbf{S}) - \omega_i g_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}))$, in such a way that the modified problem is convex[5]. In order to find a concave lower bound of $f_0(\mathbf{S})$, $g_i(\cdot)$ can be upper bounded linearly at point $\boldsymbol{\Omega}_i^{(0)} = \sum_{k\in\mathcal{U}_I, k\neq i}\mathbf{H}_i\mathbf{S}_k^{(0)}\mathbf{H}_i^H + \mathbf{I}$ as

$$\begin{aligned} & g_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i})) \\ & \quad \le g_i\left(\boldsymbol{\Omega}_i^{(0)}\right) + \operatorname{Tr}\left(\left(\boldsymbol{\Omega}_i^{(0)}\right)^{-1}\left(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}) - \boldsymbol{\Omega}_i^{(0)}\right)\right) \\ & \quad = \text{constant} + \operatorname{Tr}\left(\left(\boldsymbol{\Omega}_i^{(0)}\right)^{-1}\boldsymbol{\Omega}_i(\mathbf{S}_{-i})\right) \\ & \quad \triangleq \hat{g}_i\left(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}), \boldsymbol{\Omega}_i^{(0)}\right). \qquad (22)\end{aligned}$$

Even though problem (18) reformulated with the previous upper bound $\hat{g}_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}), \boldsymbol{\Omega}_i^{(0)})$ is convex, we want to go one step further and apply a quadratic lower bound for the left hand side of $f_0(\mathbf{S})$, i.e., $s_i(\mathbf{S})$, in a way that the overall lower bound fulfills conditions $(\mathrm{A1})-(\mathrm{A4})$ presented before in Section II-B and hence the MM method can be invoked. Note that the upper bound $\hat{g}_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}), \boldsymbol{\Omega}_i^{(0)})$ already fulfills the four conditions $(\mathrm{A1})-(\mathrm{A4})$. The idea of implementing this quadratic bound is to find a surrogate problem that is much simpler and easier to solve than the one obtained by just considering the linear bound $\hat{g}_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}), \boldsymbol{\Omega}_i^{(0)})$.[6]

We now focus attention on deriving the surrogate function for the left hand side of $f_0(\mathbf{S})$, i.e., $s_i(\mathbf{S})$. In order for the surrogate problem to be easily solved, we force the surrogate function of $s_i(\mathbf{S})$ around $\bar{\mathbf{S}}^{(0)}$ to be quadratic, where $\bar{\mathbf{S}}^{(0)} = \sum_{k\in\mathcal{U}_I}\mathbf{S}_k^{(0)}$ and $\mathbf{S}_k^{(0)}$ is the solution of the algorithm at the previous iteration. By doing this, as will be apparent later, the overall surrogate problem can be formulated as an SDP optimization problem.

Proposition 5: A valid surrogate function, $\hat{s}_i(\bar{\mathbf{S}}, \bar{\mathbf{S}}^{(0)})$, for the function $s_i(\bar{\mathbf{S}}) = \log\det\left(\mathbf{I}+\mathbf{H}_n\bar{\mathbf{S}}\mathbf{H}_n^H\right)$ that satisfies conditions $(\mathrm{A1})-(\mathrm{A4})$ is

$$\begin{aligned}\hat{s}_i(\bar{\boldsymbol{S}}, \bar{\boldsymbol{S}}^{(0)}) \triangleq & \operatorname{Tr}\left(\boldsymbol{J}_i\bar{\boldsymbol{S}}\right) + \operatorname{Tr}\left(\bar{\boldsymbol{S}}^H\boldsymbol{M}_i\bar{\boldsymbol{S}}\right) \\ & + \kappa_1, \quad \forall \bar{\boldsymbol{S}}, \bar{\boldsymbol{S}}^{(0)} \in \mathcal{S}_+^{n_T}, \qquad (23)\end{aligned}$$

with matrices $\boldsymbol{J}_i = \boldsymbol{G}_i - \bar{\boldsymbol{S}}^{(0),H}\boldsymbol{M}_i - \boldsymbol{M}_i\bar{\boldsymbol{S}}^{(0)}$, $\boldsymbol{G}_i = \boldsymbol{H}_i^H(\boldsymbol{I}+\boldsymbol{H}_i\bar{\boldsymbol{S}}^{(0)}\boldsymbol{H}_i^H)^{-1}\boldsymbol{H}_i$ and $\boldsymbol{M}_i = -\gamma_i\boldsymbol{I}$, being $\gamma_i \ge \frac{1}{2}\lambda_{\max}^2(\boldsymbol{H}_i^H\boldsymbol{H}_i)$, κ_1 contains some terms that do not depend on $\boldsymbol{S}$, and $\mathcal{S}_+^{n_T}$ denotes the set of positive semidefinite matrices.

[5]In fact, by applying the approximation, the overall objective function becomes concave.

[6]The surrogate problem obtained by just applying the bound $\hat{g}_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}), \boldsymbol{\Omega}_i^{(0)})$ will be used as benchmark. The specific mathematical details of the optimization problem and the algorithm will be described in Appendix A.

Proof: See Appendix B. ■

Let us now reformulate the optimization problem in (18) with the surrogate function $\hat{s}_i(\bar{\mathbf{S}}, \bar{\mathbf{S}}^{(0)}) - \hat{g}_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}), \mathbf{\Omega}_i^{(0)})$:

$$\mathrm{Tr}\left(\mathbf{E}_i\bar{\mathbf{S}}\right) + \mathrm{Tr}\left(\bar{\mathbf{S}}^H\mathbf{M}_i\bar{\mathbf{S}}\right) + \mathrm{Tr}\left(\mathbf{R}_i\mathbf{S}_i\right) + \kappa_2, \quad (24)$$

where $\mathbf{R}_i = \mathbf{H}_i^H(\mathbf{\Omega}_i^{(0)})^{-1}\mathbf{H}_i \in \mathbb{C}^{n_T\times n_T}$, $\mathbf{E}_i = \mathbf{J}_i - \mathbf{R}_i$, and κ_2 contains some terms that do not depend on $\mathbf{S}$. Thus, problem (18) can be reformulated as

$$\begin{aligned}\max_{\{\mathbf{S}_i\}} \quad & \sum_{i\in\mathcal{U}_I}\omega_i\left(\mathrm{Tr}\left(\mathbf{E}_i\bar{\mathbf{S}}\right) + \mathrm{Tr}\left(\bar{\mathbf{S}}^H\mathbf{M}_i\bar{\mathbf{S}}\right) + \mathrm{Tr}\left(\mathbf{R}_i\mathbf{S}_i\right)\right) \\ & -\rho\left\|\mathbf{S}_i - \mathbf{S}_i^{(0)}\right\|_F^2 \quad (25)\\ \text{s.t.} \quad & \mathbf{S}\in\mathcal{S}_1,\end{aligned}$$

where we have added a proximal quadratic term to the surrogate function in which ρ is any non-negative constant that can be tuned by the algorithm. This term provides more flexibility in the algorithm design stage and may help to speed up the convergence. By performing some mathematical manipulations, we are able to obtain the following result:

Proposition 6: The optimization problem presented in (18) can be solved based on MM method by solving recursively the following SDP problem:

$$\begin{aligned}\min_{\{\boldsymbol{S}_i\},\,\boldsymbol{s},\,t} \quad & t \quad (26)\\ \text{s.t.} \quad & C1: \begin{bmatrix} t\mathbf{I} & \tilde{\boldsymbol{C}}^{\frac{1}{2}}\boldsymbol{s}-\boldsymbol{c} \\ \left(\tilde{\boldsymbol{C}}^{\frac{1}{2}}\boldsymbol{s}-\boldsymbol{c}\right)^H & 1\end{bmatrix} \succeq 0 \\ & C2: \boldsymbol{T}_i\mathbf{s} = \mathbf{vec}\left(\boldsymbol{S}_i\right), \quad \forall i\in\mathcal{U}_I \\ & C3: \boldsymbol{S}\in\mathcal{S}_1,\end{aligned}$$

where $\mathbf{s} = [\mathrm{vec}(\boldsymbol{S}_1)^T\,\mathrm{vec}(\boldsymbol{S}_2)^T \ldots \mathrm{vec}(\boldsymbol{S}_N)^T]^T \in \mathbb{C}^{n_T n_T|\mathcal{U}_I|\times 1}$, t is a dummy variable, and $\tilde{\mathbf{C}}^{\frac{1}{2}}$, $\mathbf{T}_i$, and $\mathbf{c}$ are some constant matrices and vectors computed as shown in Appendix C. Vector $\mathbf{c}$ depends on matrix $\bar{\mathbf{S}}^{(0)}$.

Proof: See Appendix C. ■

The final algorithm is presented in Algorithm 1.

B. Approach to Solve the Sum-Based Formulation in (20)

Let us start the development by reformulating problem (20):

$$\begin{aligned}\max_{\{\mathbf{S}_i\}} \quad & \sum_{i\in\mathcal{U}_I}\omega_i\left(s_i(\mathbf{S}) - \omega_i g_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}))\right) + \sum_{i\in\mathcal{U}_I}\mathrm{Tr}(\mathbf{R}_H\mathbf{S}_i)\\ \text{s.t.} \quad & \mathbf{S}\in\mathcal{S}_2, \quad (27)\end{aligned}$$

where $\mathbf{R}_H = \sum_{j\in\mathcal{U}_E}\alpha_j\mathbf{H}_j^H\mathbf{H}_j$. The right hand side of the objective function of (27) is convex (in fact it is linear) whereas the left hand side is not convex. Let us apply the same steps that we applied before but with a slight modification. Previously in (22), we found that $g_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}))$ could be approximated by $\hat{g}_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}), \mathbf{\Omega}_i^{(0)}) = \mathrm{Tr}((\mathbf{\Omega}_i^{(0)})^{-1}\mathbf{\Omega}_i(\mathbf{S}_{-i}))$ (omitting the constant term). Now, as the objective function is different than the one from problem (18), the goal is to find a surrogate function

Algorithm 1: Algorithm for Solving Problem (18).

1: Initialize $\mathbf{S}^{(0)} \in \mathcal{S}_1$. Set $k = 0$
2: Repeat
3: Compute $\mathbf{c}$ with $\mathbf{S}^{(k)}$, given in (61)
4: Generate the $(k+1)$-th tuple $(\mathbf{S}_i^\star)_{\forall i\in\mathcal{U}_I}$ by solving the SDP in (26)
5: Set $\mathbf{S}_i^{(k+1)} = \mathbf{S}_i^\star$, $\forall i\in\mathcal{U}_I$, and set $k = k+1$
6: Until convergence is reached

for the function $s_i(\mathbf{S})$ that allows us to find efficiently a solution for the surrogate problem.

Proposition 7: A valid surrogate function, $\hat{s}_i(\mathbf{S}, \mathbf{S}^{(0)})$, for the function $s_i(\mathbf{S})$ that satisfies conditions (A1)−(A4) is

$$\hat{s}_i(\boldsymbol{S}, \boldsymbol{S}^{(0)}) \triangleq \sum_{\ell\in\mathcal{U}_I}\mathrm{Tr}\left(\boldsymbol{J}_i\boldsymbol{S}_\ell\right) + \sum_{\ell\in\mathcal{U}_I}\mathrm{Tr}\left(\boldsymbol{S}_\ell^H\boldsymbol{M}_i\boldsymbol{S}_\ell\right) + \kappa_3,$$
$$\forall \boldsymbol{S}_\ell,\ \boldsymbol{S}_\ell^{(0)} \in \mathcal{S}_+^{n_T}, \quad (28)$$

with matrices $\boldsymbol{J}_i = \boldsymbol{G}_i - \boldsymbol{S}_\ell^{(0),H}\boldsymbol{M}_i - \boldsymbol{M}_i\boldsymbol{S}_\ell^{(0)}$, $\boldsymbol{G}_i = \boldsymbol{H}_i^H(\boldsymbol{I} + \boldsymbol{H}_i\sum_{k\in\mathcal{U}_I}\boldsymbol{S}_k^{(0)}\boldsymbol{H}_i^H)^{-1}\boldsymbol{H}_i$, and $\boldsymbol{M}_i = -\xi_i\boldsymbol{I}$, with $\xi_i \geq \frac{1}{2}|\mathcal{U}_I|^2\lambda_{\max}^2(\boldsymbol{H}_i^H\boldsymbol{H}_i)$ and κ_3 containing the constant terms that do not depend on $\boldsymbol{S}$.

Proof: See Appendix D. ■

Remark 1: Note that the two surrogate functions (23) and (28) have the same form but with a difference in the quadratic term. Notice that surrogate function (28) is tighter than (23) and with cross-products. As will be shown later, this will allow us to decouple the optimization problem for each information user i and, thus, solve all problems in parallel. On the other hand, thanks to the fact that surrogate function (23) is looser than (28), a faster convergence can be obtained than if surrogate (28) were to be applied in problem (18).

Let us now reformulate problem (27) with the lower bound that we just found (omitting the constant terms):

$$\begin{aligned}\max_{\{\mathbf{S}_i\}} \quad & \sum_{i\in\mathcal{U}_I}\mathrm{Tr}\left(\check{\mathbf{J}}_i\mathbf{S}_i\right) + \sum_{i\in\mathcal{U}_I}\mathrm{Tr}\left(\mathbf{S}_i^H\check{\mathbf{M}}\mathbf{S}_i\right)\\ & -\sum_{i\in\mathcal{U}_I}\mathrm{Tr}\left(\mathbf{R}_i\sum_{\substack{k\in\mathcal{U}_I\\k\neq i}}\mathbf{S}_k\right) + \sum_{i\in\mathcal{U}_I}\mathrm{Tr}(\mathbf{R}_H\mathbf{S}_i) \quad (29)\\ \text{s.t.} \quad & \mathbf{S}\in\mathcal{S}_2,\end{aligned}$$

where $\check{\mathbf{J}}_i = \check{\mathbf{G}} - \mathbf{S}_i^{(0),H}\check{\mathbf{M}} - \check{\mathbf{M}}\mathbf{S}_i^{(0)}$, with $\check{\mathbf{M}} = \sum_{k\in\mathcal{U}_I}\omega_k\mathbf{M}_k$ and $\check{\mathbf{G}} = \sum_{k\in\mathcal{U}_I}\omega_k\mathbf{G}_k$. Note that we have arranged the indices to make the notation easier to follow and consistent with the original notation. We can further simplify the objective function by grouping terms considering that matrix $\check{\mathbf{M}}$ is diagonal, i.e., $\check{\mathbf{M}} = -\beta\mathbf{I}$, being $\beta = \frac{1}{2}|\mathcal{U}_I|^2\sum_{k\in\mathcal{U}_I}\omega_k\lambda_{\max}^2(\mathbf{H}_k^H\mathbf{H}_k)$:

$$\begin{aligned}\min_{\{\mathbf{S}_i\}} \quad & \beta\sum_{i\in\mathcal{U}_I}\mathrm{Tr}\left(\mathbf{S}_i^H\mathbf{S}_i\right) - \sum_{i\in\mathcal{U}_I}\mathrm{Tr}\left(\mathbf{F}_i\mathbf{S}_i\right) \quad (30)\\ \text{s.t.} \quad & \mathbf{S}\in\mathcal{S}_2,\end{aligned}$$

where

$$\mathbf{F}_i = \check{\mathbf{J}}_i - \sum_{\substack{k\in\mathcal{U}_I \\ k\neq i}} \mathbf{R}_k + \mathbf{R}_H. \tag{31}$$

Note that we have changed the sign of the objective and reformulated the problem as a minimization one. The idea is to find a closed-form expression for the optimum covariance matrices $\{\mathbf{S}_i\}$. If we dualize constraint $C1$ and form a partial Lagrangian, we obtain the following optimization problem:

$$\begin{aligned} \min_{\{\mathbf{S}_i\}} \quad & \beta \sum_{i\in\mathcal{U}_I} \mathrm{Tr}\left(\mathbf{S}_i^H \mathbf{S}_i\right) - \sum_{i\in\mathcal{U}_I} \mathrm{Tr}\left(\mathbf{W}_i(\mu)\mathbf{S}_i\right) \\ \text{s.t.} \quad & \mathbf{S}_i \succeq 0, \quad \forall i \in \mathcal{U}_I, \end{aligned} \tag{32}$$

where $\mathbf{W}_i(\mu) = \mathbf{F}_i - \mu\mathbf{I}$, for $\mu \geq 0$ the Lagrange multiplier associated with constraint $C1$ of problem (27). The previous problem is clearly separable for each user i. Thus, for each information user, problem (32) is equivalent to solving the following projection problem:

$$\begin{aligned} \min_{\mathbf{S}_i} \quad & \left\| \sqrt{\beta}\mathbf{S}_i - \check{\mathbf{W}}_i(\mu) \right\|_F \\ \text{s.t.} \quad & \mathbf{S}_i \succeq 0, \end{aligned} \tag{33}$$

where $\check{\mathbf{W}}_i(\mu) = \frac{1}{2\sqrt{\beta}}\mathbf{W}_i(\mu) = \frac{1}{2\sqrt{\beta}}(\mathbf{F}_i - \mu\mathbf{I})$. The previous result is very nice as the solution of (33) is simple and elegant, thanks to the fact that problem (33) is a projection onto the semidefinite cone and has a closed-form solution [28]. Let the eigenvalue decomposition (EVD) of matrix $\mathbf{F}_i$ be $\mathbf{F}_i = \mathbf{U}_{F_i}\mathbf{\Lambda}_{F_i}\mathbf{U}_{F_i}^H$. The expression of $\mathbf{S}_i^\star(\mu)$ is, thus, given by

$$\mathbf{S}_i^\star(\mu) = \frac{1}{\sqrt{\beta}}[\check{\mathbf{W}}_i(\mu)]^+ = \frac{1}{2\beta}\mathbf{U}_{F_i}^H[\mathbf{\Lambda}_{F_i} - \mu\mathbf{I}]^+\mathbf{U}_{F_i}, \ \forall i \in \mathcal{U}_I, \tag{34}$$

where $\lambda_k([\mathbf{X}]^+) = \min(0, \lambda_k(\mathbf{X}))$, with $\lambda_k(\mathbf{X})$ the k-th eigenvalue of matrix $\mathbf{X}$. Now it remains to compute the optimal Lagrange multiplier μ. This can be found by means of the simple bisection method fulfilling $\sum_{i\in\mathcal{U}_I} \mathrm{Tr}\left([\mathbf{\Lambda}_{F_i} - \mu\mathbf{I}]^+\right) = 2\beta P_T$. It turns out that, at each inner iteration, we need to compute a single EVD per information user, that is, the EVD of $\mathbf{F}_i$, and then a few iterations are needed to find the optimal multiplier μ (using for example the bisection method in step 5 of Algorithm 2). Note that the surrogate problem can be solved straightforwardly with the previous steps. The final algorithm is presented in Algorithm 2.

C. Approaches Used as Benchmarks for Performance Comparison

In this section, we propose some benchmark algorithms that will be used in the simulations section to assess the performance of the MM approaches proposed in the previous subsections. These benchmarks have been derived from previous works and are the following:

- Gradient-based algorithms based on [29, Sec. 7] applied directly to the nonconvex problems (18) and (20). The expressions of the gradients are not presented here due to space limitations but are developed in the detail by the same authors in [16].
- MM approaches considering just the linear approximation presented in (22), i.e., $\hat{g}_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}), \mathbf{\Omega}_i^{(0)})$, applied to problems (18) and (20). The specific optimization problems and algorithms (which were briefly addressed in [16]) can be found in Appendix A.
- Optimization of the sum-rate based on its relation with the MSE. This relation was exploited in [15] to deduce a block-based alternating optimization algorithm; however, no harvesting constraints were considered. The inclusion of harvesting constraints was addressed in [30] by means of an iterative method in which those constraints were simplified through successive linear approximations. The simulations section (Section V) presents as a benchmark the method developed in [30] but adapted to a multiuser system following the same approach as in [15].

Algorithm 2: Algorithm for Solving Problem (20).

1: Initialize $\mathbf{S}^{(0)} \in \mathcal{S}_2$. Set $k = 0$
2: Repeat
3: Compute $\mathbf{F}_i$ with matrix $\mathbf{S}_i^{(k)}$, $\forall i \in \mathcal{U}_I$, given in (31)
4: Compute EVD of $\mathbf{F}_i = \mathbf{U}_{F_i}\mathbf{\Lambda}_{F_i}\mathbf{U}_{F_i}^H, \quad \forall i \in \mathcal{U}_I$
5: Compute $\mu^\star$ such that $\sum_{i\in\mathcal{U}_I} \mathrm{Tr}\left([\mathbf{\Lambda}_{F_i} - \mu^\star\mathbf{I}]^+\right) = 2\beta P_T$
6: Compute $\mathbf{S}_i^\star(\mu^\star) = \frac{1}{2\beta}[\mathbf{F}_i - \mu^\star\mathbf{I}]^+, \quad \forall i \in \mathcal{U}_I$
7: Set $\mathbf{S}_i^{(k+1)} = \mathbf{S}_i^\star(\mu^\star), \forall i \in \mathcal{U}_I$, and set $k = k + 1$
8: Until convergence is reached

V. Numerical Evaluation

In this section, we evaluate the performance of the previous algorithms. In the first part of this section, we present some convergence and computational time results. For the simulations, we consider a system composed of 1 transmitter with 6 antennas along with 3 information users and 3 harvesting users with 2 antennas each. In the second part of the section, we show the performance of the proposed methods compared to the classical BD approach. In this case, for ease of presenting the information, we assume a system composed of 1 transmitter with 4 antennas, and 2 information users and 2 harvesting users with 2 antennas each. The simulation parameters common to both scenarios are the following. The maximum radiated power is $P_T = 1$ W. The channel matrices are generated randomly with i.i.d. entries distributed according to $\mathcal{CN}(0, 1)$. The weights ω_i are set to 1.

A. Convergence Evaluation

In this section, we evaluate the convergence behavior and the computational time of the methods presented in Sections IV-A and IV-B and the benchmark approach presented in Appendix A. The benchmark method for problem (20) presented in Appendix A will not be evaluated as it is clearly worse[7]

[7]However, it was included in the paper for the sake of completeness.

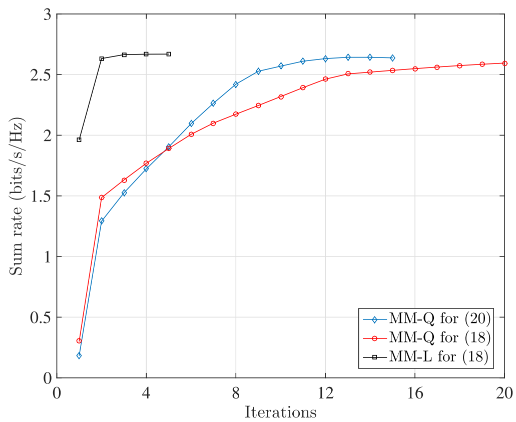

Fig. 1. Convergence of the system sum rate vs number of iterations for three different approaches.

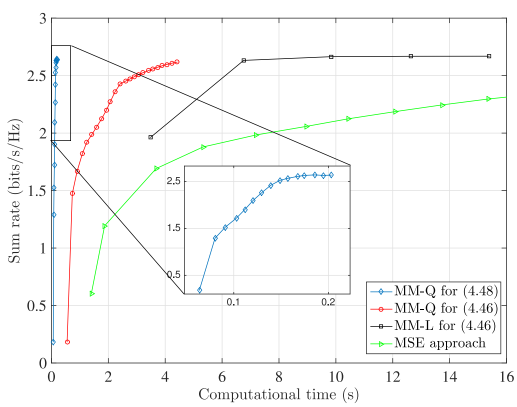

Fig. 2. Convergence of the system sum rate vs computational time for three different approaches.

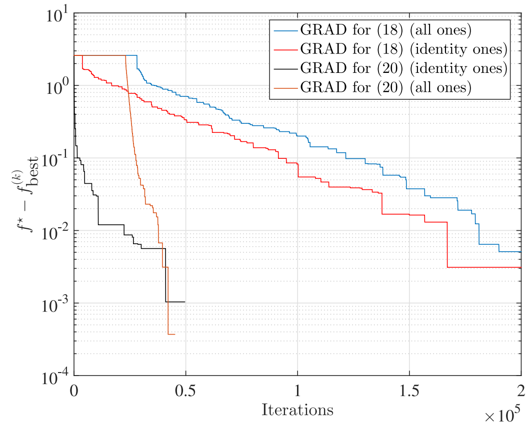

Fig. 3. Convergence of the system sum rate vs iterations for a gradient approach for constrained optimization.

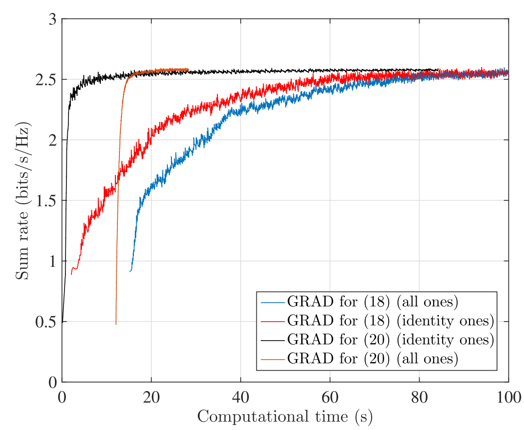

Fig. 4. Convergence of the system sum rate vs computational time for a gradient approach for constrained optimization.

than the one presented in Section IV-B. In the figures, the legend is interpreted as follows: 'MM-L for (18)' refers to the method developed in Appendix A for problem (18), 'MM-Q for (18)' refers to the method in Section IV-A, and 'MM-Q for (20)' refers to the method in Section IV-B. In order to compare all methods, we set the values of α_j and the values of Q_j so that the same system sum rate is achieved. These values are: $\boldsymbol{\alpha} = [1, 5, 10]$, and $\mathbf{Q} = [3.8, 7.2, 6.4]$ power units. Software package CVX is used to solve problem (35) [31], and SeDuMi solver is used to solve problem (26) [32].

Fig. 1 presents the sum rate convergence as a function of iterations. The three approaches converge to the same sum rate value but require a different number of iterations. In fact, the required number of iterations depends on how well the surrogate function approximates the original function. Note that the surrogate function used in the 'MM-L for (18)' approach is the one that best approximates the objective function and, thus, fewer iterations are needed.

Fig. 2 shows the computational time required by the three previous methods and the benchmark based on the 'MSE approach' [30]. We see that the 'MM-Q for (20)' method converges much faster than the other two approaches, as expected. The 'MM-Q for (18)' approach requires more iterations than the 'MM-L for (18)' approach but each iteration is solved faster since a specific algorithm can be employed to solve the convex optimization problem. Hence, the 'MM-Q for (18)' algorithm is the best option. Additionally, we clearly see that the proposed MM method is much faster than the method based on the MSE.

For the sake of comparison and completeness, we also show in Figs. 3 and 4 the convergence and the computational time of a gradient-like benchmark approach. The plot legend reads as follows: 'GRAD for (18)' and 'GRAD for (20)' refers to a gradient approach applied to problems (18) and (20), respectively. 'all ones' and 'identity' mean that covariance matrices are initialized using an all ones matrix and the identity matrix, respectively. Results show that the proposed MM approaches are one to two orders of magnitude faster than the gradient-based methods.

B. *Performance Evaluation*

In this section, we evaluate the performance of the MM approach as compared to the classical BD strategy considered in the literature (see, for example, [6], [33]). In order to show how harvesting users at different distances (and, hence, path loss) affect the performance, we have generated the channel matrices in a way that the there is a factor of 2 in the Frobenius norm of those matrices. We would like to emphasize that, as the noise and channels are normalized, we will refer to the powers harvested by the receivers in terms of power units instead of Watts.

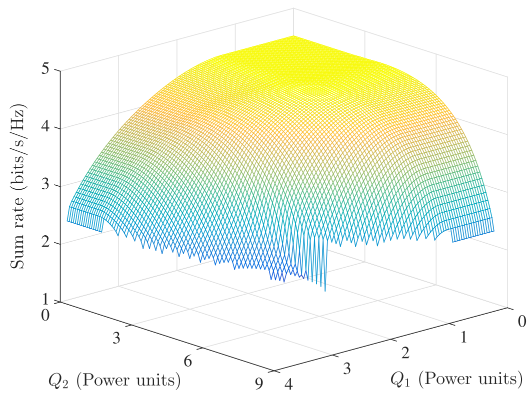

Fig. 5. Rate-power surface for the MM method.

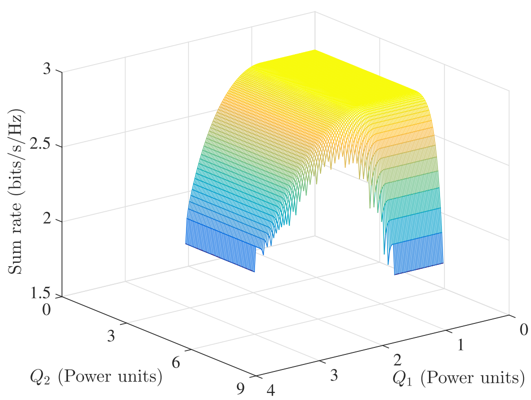

Fig. 6. Rate-power surface for the BD method.

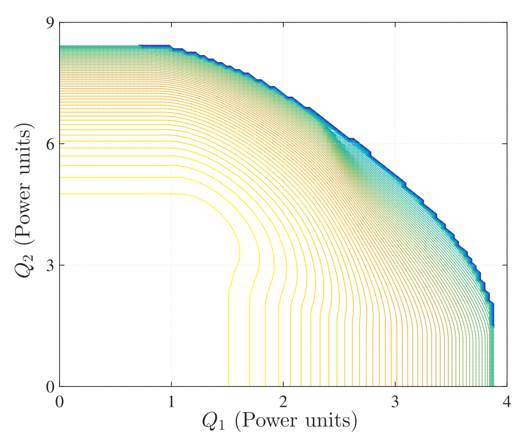

Fig. 7. Contour of rate-power surface for the MM method.

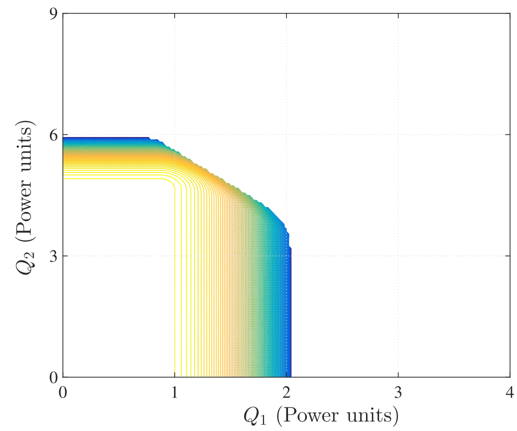

Fig. 8. Contour of rate-power surface for the BD method.

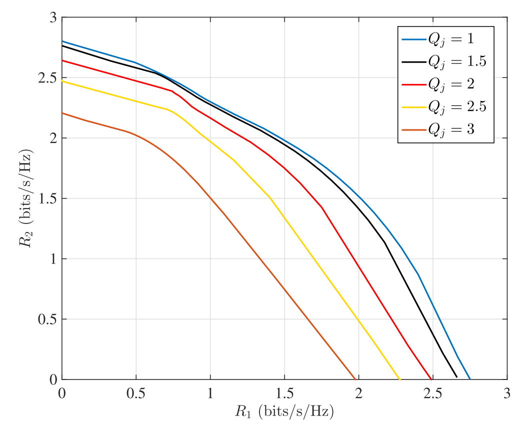

Fig. 9. Rate region for different values of Q_j (in power units).

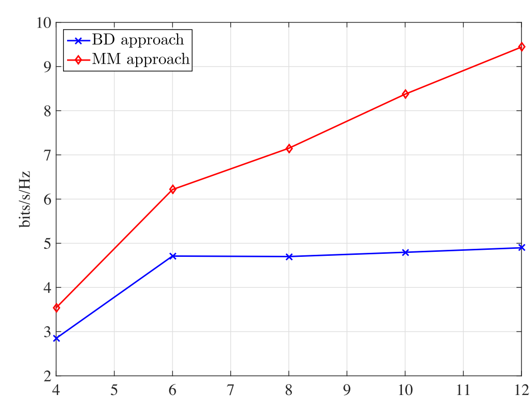

Fig. 10. System sum rate as a function of the number of transmit antennas.

Figs. 5 and 6 show the rate-power surface, that is, the multidimensional trade-off between the system sum rate and the powers to be collected by harvesting users (see [6] for a formal definition of the rate-power surface). As we see, the MM approach outperforms the BD strategy in both terms, sum rate and harvested power. The maximum system sum rate obtained with the MM approach when Q_1 and Q_2 are set to 0 is 4.5 bit/s/Hz, whereas the sum rate obtained with the BD approach is 2.75 bit/s/Hz. The rate-power surfaces are generated by varying the values of $\{Q_j\}$ in problem (18) or, equivalently, by varying the values of $\{\alpha_j\}$ in problem (20). A way to reduce the computational complexity associated with the generation of the rate-power surface is to use as an initialization point the solution that was obtained for the previous values of $\{Q_j\}$ or $\{\alpha_j\}$ to generate the new value of the curve [34]. Note, however, that the whole rate-power surface need not be generated for each transmission as it is just the representation of the existing rate-power tradeoff.

In order to clearly see the benefits in terms of collected power, Figs. 7 and 8 show the contour plots of the previous 3D plots. We observe that users in the MM approach collect roughly 50% more power than the power collected by users when applying the BD strategy.

Finally, Fig. 9 presents the rate-region of the MM approach for different values of $\{Q_j\}$. The same value of Q_j is set to the

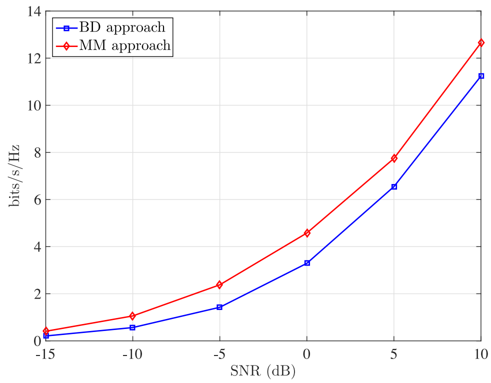

Fig. 11. System sum rate as a function of SNR.

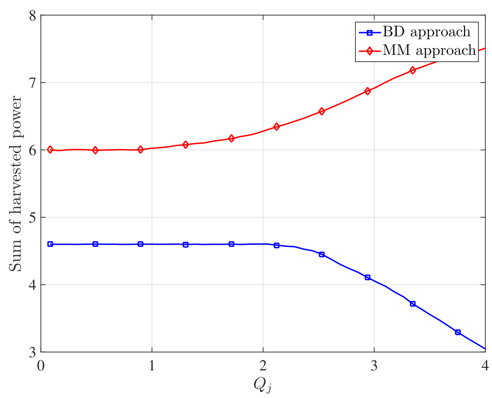

Fig. 12. Power harvested by all users as a function of the minimum powers to be harvested Q_j.

two harvesting users. In this case, we vary the values of ω_i to achieve the whole contour of the rate regions. We observe that, the larger the harvesting constraints, the smaller the rate-region, as expected. However, the relation between the harvesting constraints and the rate-region is not linear. As the harvesting constraints increase, a small change in the $\{Q_j\}$ produces a large reduction of the rate-region. This is because the 3D rate-power surfaces presented before are not planes. In the following, we analyze the performance of both approaches, namely MM and BD, as a function of several system parameters to obtain valuable insights into the proposed scheme. First, in Fig. 10, we show the dependence of the system sum rate with respect to the number of transmit antennas for both methods. As we see, the proposed MM method outperforms the BD scheme for all antenna configurations, specially for larger number of transmit antennas, where we see that the sum rate of the BD approach tends to saturate whereas the sum rate of the MM method increases quite fast. In Fig. 11, we plot the system sum rate as a function of the SNR, where the SNR is defined as P_T/σ^2. In this case, the difference between the two methods is not that significant, but there is still an improvement of the sum rate obtained with the MM method with respect to the BD method.

Finally, Fig. 12 shows the sum of the actual powers harvested by all users in the system as a function of the minimum powers to be harvested, Q_j, introduced through constraints. If we focus on the MM method, for values of Q_j smaller than 2, the harvesting constraints are not active since the value of sum power obtained with no constraints is 6 power units and there are 3 harvesting users (all of them configured with the same value of Q_j). For larger values of Q_j, the harvesting constraints start to activate. In some cases, specially when the values of Q_j are high, the optimization problem may turn out to be not feasible for some realizations and the sum power obtained is lower than the one expected (since the obtained sum power is set to 0 in the realizations in which the problem results to be non-feasible). For example, for $Q_j = 3$ a sum power of 9 units should be obtained instead of 7. If we have a look at the BD method, we see that the system behaves even worse. For larger values of Q_j the overall sum power is lower than the sum power obtained for small values of Q_j. This phenomenon is due to the fact the problem corresponding to BD turns out to be non-feasible more frequently than in the case of not applying the BD constraints. Hence, from this figure we conclude that the MM method is superior to the BD approach also in terms of actual harvested powers.

VI. Conclusion

We have presented a method to solve the difficult nonconvex problem that arises in multiuser multi-stream broadcast MIMO SWIPT networks. We formulated the general SWIPT problem as a multi-objective optimization problem, in which rates and harvested powers were to be optimized simultaneously. Then, we proposed two different formulations to obtain solutions of the general multi-objective optimization problem depending on the desired level of control of the power to be harvested. In the first approach, the transmitter was able to control the specific amount of power to be harvested by each user whereas in the second approach only the proportions of power to be harvested among the different users could be controlled. Both (nonconvex) formulations were solved based on the MM approach. We derived a convex approximation for two nonconvex objectives and developed two different algorithms. Simulation results showed that the proposed methods outperform the classical BD, in terms of both system sum rate and power collected by users, by a factor of approximately 50%. Moreover, the computational time needed to achieve convergence was shown to be really low for the approach in which the transmitter could only control the proportion of powers to be harvested (around two orders of magnitude lower than a gradient-like approach).

There are some research lines that can be considered to further extend the work presented in this paper. Firstly, nonlinear energy harvesting constraints could be considered as they model nonlinearities found in practical energy harvesting receivers. Having nonlinear harvesting constraints increases the complexity of the overall solution and finding efficient algorithms is a challenge. Secondly, it would be interesting to consider the case of having imperfect CSI at the transmitter.

Appendix A
Benchmark Formulations and Algorithms

In this appendix, we are going to describe the benchmarks based on the works in [25], [26], and [27]. We start with the benchmark for problem (18).

Algorithm 3: Algorithm for Solving Problem (18).

1: Initialize $\mathbf{S}^{(0)} \in \mathcal{S}_1$. Set $k = 0$
2: Repeat
3: Generate the $(k+1)$-th tuple $(\mathbf{S}_i^{\star})_{\forall i \in \mathcal{U}_I}$ by solving (35)
4: Set $\mathbf{S}_i^{(k+1)} = \mathbf{S}_i^{\star}$, $\forall i \in \mathcal{U}_I$, and set $k = k+1$
5: Until convergence is reached

Note that the upper bound $\hat{g}_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}), \boldsymbol{\Omega}_i^{(0)})$ can be used to build a lower bound of $f_0(\bar{\mathbf{S}})$ that fulfills the four conditions $(\mathrm{A1})-(\mathrm{A4})$ presented before in Section II-B.

By applying a successive approximation of $f_0(\cdot)$ through the application of the previous surrogate function, i.e., $\hat{f}_0(\mathbf{S}, \mathbf{S}^{(k)}) = \sum_{i \in \mathcal{U}_I} \omega_i s_i(\mathbf{S}) - \omega_i \hat{g}_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}), \boldsymbol{\Omega}_i^{(k)}) - \rho \|\mathbf{S}_i - \mathbf{S}_i^{(k)}\|_F^2$, where $\mathbf{S}^{(k)} \triangleq (\mathbf{S}_i^{(k)})_{\forall i \in \mathcal{U}_I}$, for different evaluation points, we obtain an iterative algorithm based on the MM approach that converges to a stationary point (or local optimum) of the original problem (18). Note that we have considered a proximal-like term. Given this, the convex optimization problem to solve is

$$\begin{aligned} \max_{\{\mathbf{S}_i\}} \quad & \sum_{i \in \mathcal{U}_I} \omega_i s_i(\mathbf{S}) - \omega_i \hat{g}_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}), \boldsymbol{\Omega}_i^{(k)}) - \rho \left\| \mathbf{S}_i - \mathbf{S}_i^{(k)} \right\|_F^2 \\ \text{s.t.} \quad & \mathbf{S} \in \mathcal{S}_1. \end{aligned} \tag{35}$$

We must proceed iteratively until convergence is reached. The procedure is presented in Algorithm 3.

Let us now continue with the benchmark for problem (20). If we apply the bound from (22), i.e., $\hat{g}_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}), \boldsymbol{\Omega}_i^{(0)})$, problem (20) can be solved by solving consecutively the following problem:

$$\begin{aligned} \max_{\{\mathbf{S}_i\}} \quad & \sum_{i \in \mathcal{U}_I} \omega_i s_i(\mathbf{S}) - \omega_i \hat{g}_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}), \boldsymbol{\Omega}_i^{(k)}) + \mathrm{Tr}(\mathbf{R}_H \mathbf{S}_i) \\ & - \rho \left\| \mathbf{S}_i - \mathbf{S}_i^{(k)} \right\|_F^2 \\ \text{s.t.} \quad & \mathbf{S} \in \mathcal{S}_2. \end{aligned} \tag{36}$$

As problem (36) is convex, the MM method can be invoked to obtain a local optimum of problem (20), following the same procedure as we did before for problem (35).

Appendix B
Proof of Proposition 5

The proposed quadratic surrogate function of $s_i(\bar{\mathbf{S}})$ has the following form:

$$\begin{aligned} \hat{s}_i(\bar{\mathbf{S}}, \bar{\mathbf{S}}^{(0)}) \triangleq\ & \log\det\left(\mathbf{I} + \mathbf{H}_i \bar{\mathbf{S}}^{(0)} \mathbf{H}_i^H\right) \\ & + \mathrm{Re}\Big\{ \mathrm{Tr}\left(\mathbf{G}_i \left(\bar{\mathbf{S}} - \bar{\mathbf{S}}^{(0)}\right)\right) \Big\} \\ & + \mathrm{Tr}\left(\left(\bar{\mathbf{S}} - \bar{\mathbf{S}}^{(0)}\right)^H \mathbf{M}_i \left(\bar{\mathbf{S}} - \bar{\mathbf{S}}^{(0)}\right)\right) \\ \leq\ & \log\det\left(\mathbf{I} + \mathbf{H}_i \bar{\mathbf{S}} \mathbf{H}_i^H\right), \quad \forall \bar{\mathbf{S}},\, \bar{\mathbf{S}}^{(0)} \in \mathcal{S}_+^{n_T}, \end{aligned} \tag{37}$$

where matrices $\mathbf{G}_i \in \mathbb{C}^{n_T \times n_T}$ and $\mathbf{M}_i \in \mathbb{C}^{n_T \times n_T}$ need to be found such that conditions (A1) through (A4) are satisfied, and $\mathrm{Re}\{x\}$ denotes the real part of x. Note that (A1) and (A4) are already satisfied. Only (A2) and (A3) must be ensured.

Let us start by proving condition (A3). Let $\bar{\mathbf{S}}^{(0)}$ and $\bar{\mathbf{S}}^{(1)}$ be two positive semidefinite matrices, i.e, $\bar{\mathbf{S}}^{(0)}, \bar{\mathbf{S}}^{(1)} \in \mathcal{S}_+^{n_T}$. Then, the directional derivative of the surrogate function $\hat{s}_i(\bar{\mathbf{S}}, \bar{\mathbf{S}}^{(0)})$ in (37) at $\bar{\mathbf{S}}^{(0)}$ with direction $\bar{\mathbf{S}}^{(1)} - \bar{\mathbf{S}}^{(0)}$ is given by:

$$\mathrm{Re}\Big\{ \mathrm{Tr}\left(\mathbf{G}_i \left(\bar{\mathbf{S}}^{(1)} - \bar{\mathbf{S}}^{(0)}\right)\right) \Big\}. \tag{38}$$

Now, let us compute the directional derivative of the term $\log\det\left(\mathbf{I} + \mathbf{H}_i \bar{\mathbf{S}} \mathbf{H}_i^H\right)$:

$$\mathrm{Tr}\left(\mathbf{H}_i^H \left(\mathbf{I} + \mathbf{H}_i \bar{\mathbf{S}}^{(0)} \mathbf{H}_i^H\right)^{-1} \mathbf{H}_i \left(\bar{\mathbf{S}}^{(1)} - \bar{\mathbf{S}}^{(0)}\right)\right), \tag{39}$$

where we have used $\mathrm{d}\log\det(\mathbf{X}) = \mathrm{Tr}(\mathbf{X}^{-1}\mathrm{d}\mathbf{X})$ [35]. Hence, by applying condition (A3), the two directional derivatives (38) and (39) must be equal, from which we are able to identify matrix $\mathbf{G}_i$ as

$$\mathbf{G}_i = \mathbf{H}_i^H \left(\mathbf{I} + \mathbf{H}_i \bar{\mathbf{S}}^{(0)} \mathbf{H}_i^H\right)^{-1} \mathbf{H}_i, \qquad \mathbf{G}_i = \mathbf{G}_i^H. \tag{40}$$

equation (41), (42), (43), (44) as shown at the bottom of next page.

Note that as matrix $\mathbf{G}_i$ is hermitian, the real operator is no longer needed since the trace of the product of two hermitian matrices is real. In order to prove condition (A2), it suffices to show that for each linear cut in any direction, the surrogate function is a lower bound. Let $\bar{\mathbf{S}} = \bar{\mathbf{S}}^{(0)} + \mu\left(\bar{\mathbf{S}}^{(1)} - \bar{\mathbf{S}}^{(0)}\right)$, $\forall \mu \in [0,1]$. Then, it suffices to show (41). Now, a sufficient condition for (41) is that the second derivative of the left hand side of (41) is lower than or equal to the second derivative of the right hand side of (41) for any $\mu \in [0,1]$ and any $\bar{\mathbf{S}}^{(1)}, \bar{\mathbf{S}}^{(0)} \in \mathcal{S}_+^{n_T}$, which is formulated in (42).[8]

Let us compute the second derivative of the right hand side of (42). The first derivative is given by (43) and the second derivative is given by (44), where we have used the identity $\mathrm{d}\mathbf{X}^{-1} = -\mathbf{X}^{-1}\mathrm{d}\mathbf{X}\mathbf{X}^{-1}$ [35] and matrix $\mathbf{A}_i \in \mathbb{C}^{n_{R_i} \times n_{R_i}}$ is defined as $\mathbf{A}_i = \mathbf{I} + \mathbf{H}_i(\bar{\mathbf{S}}^{(0)} + \mu\left(\bar{\mathbf{S}}^{(1)} - \bar{\mathbf{S}}^{(0)}\right))\mathbf{H}_i^H$.

We need to manipulate the previous expressions. To this end, let us define matrix $\mathbf{P}_i = \mathbf{H}_i^H \mathbf{A}_i^{-1} \mathbf{H}_i \in \mathbb{C}^{n_T \times n_T}$ and let us vectorize the result found in (44):

$$\begin{aligned} & \mathrm{Tr}\left(\mathbf{P}_i \left(\bar{\mathbf{S}}^{(1)} - \bar{\mathbf{S}}^{(0)}\right) \mathbf{P}_i \left(\bar{\mathbf{S}}^{(1)} - \bar{\mathbf{S}}^{(0)}\right)\right) \\ = & \ \mathrm{vec}\left(\left(\bar{\mathbf{S}}^{(1)} - \bar{\mathbf{S}}^{(0)}\right)^T\right)^T \left(\mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i\right) \mathrm{vec}\left(\bar{\mathbf{S}}^{(1)} - \bar{\mathbf{S}}^{(0)}\right), \end{aligned} \tag{45}$$

where we have used the following properties: $\mathrm{Tr}(\mathbf{AB}) = \mathrm{vec}(\mathbf{A}^T)^T \mathrm{vec}(\mathbf{B})$, $\quad \mathrm{vec}(\mathbf{AB})^T = \mathrm{vec}(\mathbf{A})^T(\mathbf{I} \otimes$

[8] Expression (42) is equivalent to finding a constant (left hand side of (42)) such that this constant is lower than or equal to the second derivative of the logdet function (right hand side of (42)). Then, If we take this inequality and apply a definite integration at both sides twice between $\mu = 0$ and a generic $\mu \in [0,1]$, then the inequality still holds. In fact equation (41) results from applying the previous methodology. This proves that expression (42) is a sufficient condition for (41).

$\mathbf{B}$), $\text{vec}(\mathbf{AB}) = (\mathbf{I}\otimes\mathbf{A})\text{vec}(\mathbf{B})$, and $(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D}) = (\mathbf{AC})\otimes(\mathbf{BD})$. Let us now vectorize the left hand side of (42):

$$2\ \text{Tr}\left(\left(\bar{\mathbf{S}}-\bar{\mathbf{S}}^{(0)}\right)^H \mathbf{M}_i\left(\bar{\mathbf{S}}-\bar{\mathbf{S}}^{(0)}\right)\right) = 2\text{vec}\left(\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)^T\right)^T (\mathbf{I}\otimes\mathbf{M}_i)\text{vec}\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right), \tag{46}$$

where in (46) we have used the fact that $\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}$ is hermitian and $\text{Tr}(\mathbf{ABC}) = \text{vec}(\mathbf{A}^T)^T(\mathbf{I}\otimes\mathbf{B})\text{vec}(\mathbf{C})$. Finally, we end up with the relation from forcing that (46) must be lower than or equal to (45). This relation can be expressed as given by (47), shown at the bottom of the page.

A sufficient condition for expression (47) is

$$(\mathbf{I}\otimes\mathbf{M}_i)+\frac{1}{2}\left(\mathbf{I}\otimes\mathbf{P}_i^T\mathbf{P}_i\right) = \mathbf{I}\otimes\left(\mathbf{M}_i+\frac{1}{2}\mathbf{P}_i^T\mathbf{P}_i\right)\preceq 0, \tag{48}$$

which means that

$$\mathbf{M}_i+\frac{1}{2}\mathbf{P}_i^T\mathbf{P}_i\preceq 0. \tag{49}$$

Now, if we set $\mathbf{M}_i=\alpha\mathbf{I}$ (note that this is a particular simple solution), we have that

$$\alpha\leq-\frac{1}{2}\lambda_{\max}\left(\mathbf{P}_i^T\mathbf{P}_i\right), \tag{50}$$

where $\lambda_{\max}(\mathbf{X})$ is the maximum eigenvalue of matrix $\mathbf{X}$. Now, let us introduce the following result:

Theorem 1 ([36]): Let $\mathbf{A},\ \mathbf{B}\in\mathbb{C}^{n\times n}$, assume that $\mathbf{A}$ is positive definite, and assume that $\mathbf{B}$ is positive definite. Let $\lambda_i(\mathbf{A})$ be the i-th eigenvalue of matrix $\mathbf{A}$ such that $\lambda_1(\mathbf{A})\geq\lambda_2(\mathbf{A})\geq\ldots\geq\lambda_n(\mathbf{A})$. Then, for all $i,j,k\in\{1,\ldots,n\}$ such that $j+k\leq i+1$,

$$\lambda_i(\mathbf{AB})\leq\lambda_j(\mathbf{A})\lambda_k(\mathbf{B}). \tag{51}$$

In particular, for all $i=1,\ldots,n$,

$$\lambda_i(\mathbf{A})\lambda_n(\mathbf{B})\leq\lambda_i(\mathbf{AB})\leq\lambda_i(\mathbf{A})\lambda_1(\mathbf{B}). \tag{52}$$

Thanks to the previous result, $\alpha\leq-\frac{1}{2}\lambda^2_{\max}(\mathbf{P}_i)$. Now, let the singular value decomposition of $\mathbf{H}_i$ be $\mathbf{H}_i=\mathbf{U}_i\mathbf{\Sigma}_i\mathbf{V}_i^H$. From this, we can upper bound $\lambda_{\max}(\mathbf{P}_i)=\lambda_{\max}\left(\mathbf{H}_i^H\mathbf{A}_i^{-1}\mathbf{H}_i\right)=\lambda_{\max}\left(\mathbf{\Sigma}_i\ \mathbf{V}_i^H\ \mathbf{A}_i^{-1}\mathbf{V}_i\ \mathbf{\Sigma}_i\right)\leq\sigma^2_{\max}(\mathbf{H}_i)\,\lambda^{-1}_{\min}(\mathbf{A}_i)$, where $\sigma_{\max}(\mathbf{X})$ is the maximum singular value of matrix $\mathbf{X}$. Because matrix $\mathbf{A}$ is positive definite with $\lambda_{\min}(\mathbf{A}_i)\geq 1$, we can conclude that

$$\alpha\leq-\frac{1}{2}\sigma^4_{\max}(\mathbf{H}_i), \tag{53}$$

and thus, a possible matrix $\mathbf{M}_i$ satisfying conditions (A1)–(A4) is finally

$$\mathbf{M}_i=-\frac{1}{2}\sigma^4_{\max}(\mathbf{H}_i)\mathbf{I}=-\frac{1}{2}\lambda^2_{\max}(\mathbf{H}_i^H\mathbf{H}_i)\mathbf{I}. \tag{54}$$

APPENDIX C
PROOF OF PROPOSITION 6

Let us start by vectorizing the surrogate function in (24):

$$\begin{aligned}\hat{R}_i(\mathbf{S},\mathbf{S}^{(0)}) &= \hat{s}_i(\bar{\mathbf{S}},\bar{\mathbf{S}}^{(0)})-\hat{g}_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}),\mathbf{\Omega}_i^{(0)})\\ &= \text{vec}\left(\bar{\mathbf{S}}^T\right)^T(\mathbf{I}\otimes\mathbf{M}_i)\,\text{vec}\left(\bar{\mathbf{S}}\right)+\mathbf{e}_i^T\text{vec}\left(\bar{\mathbf{S}}\right)\\ &\quad+\mathbf{r}_i^T\text{vec}\left(\mathbf{S}_i\right)+\kappa_2,\end{aligned} \tag{55}$$

$$\begin{aligned}&\log\det\left(\mathbf{I}+\mathbf{H}_i\bar{\mathbf{S}}^{(0)}\mathbf{H}_i^H\right)+\mu\,\text{Tr}\left(\mathbf{G}_i\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\right)+\mu^2\,\text{Tr}\left(\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)^H\mathbf{M}_i\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\right)\\ &\leq\log\det\left(\mathbf{I}+\mathbf{H}_i\left(\bar{\mathbf{S}}^{(0)}+\mu\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\right)\mathbf{H}_i^H\right),\quad\forall\bar{\mathbf{S}}^{(1)},\,\bar{\mathbf{S}}^{(0)}\in\mathcal{S}_+^{n_T},\,\forall\mu\in[0,1].\end{aligned} \tag{41}$$

$$2\,\text{Tr}\left(\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)^H\mathbf{M}_i\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\right)\leq\frac{\partial^2}{\partial\mu^2}\log\det\left(\mathbf{I}+\mathbf{H}_i\left(\bar{\mathbf{S}}^{(0)}+\mu\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\right)\mathbf{H}_i^H\right)\Bigg|_{\forall\bar{\mathbf{S}}^{(1)},\bar{\mathbf{S}}^{(0)}\in\mathcal{S}_+^{n_T},\,\forall\mu\in[0,1]}. \tag{42}$$

$$\begin{aligned}&\frac{\partial}{\partial\mu}\log\det\left(\mathbf{I}\ +\ \mathbf{H}_i\left(\bar{\mathbf{S}}^{(0)}+\mu\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\right)\mathbf{H}_i^H\right)\\ &\quad=\text{Tr}\left(\left(\mathbf{I}+\mathbf{H}_i\left(\bar{\mathbf{S}}^{(0)}+\mu\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\right)\mathbf{H}_i^H\right)^{-1}\mathbf{H}_i\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\mathbf{H}_i^H\right),\end{aligned} \tag{43}$$

$$\frac{\partial^2}{\partial\mu^2}\log\det\left(\mathbf{I}+\mathbf{H}_i\left(\bar{\mathbf{S}}^{(0)}+\mu\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\right)\mathbf{H}_i^H\right)=-\,\text{Tr}\left(\mathbf{A}_i^{-1}\mathbf{H}_i\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\mathbf{H}_i^H\mathbf{A}_i^{-1}\mathbf{H}_i\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\mathbf{H}_i^H\right), \tag{44}$$

$$2\text{vec}\left(\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)^T\right)^T\left[(\mathbf{I}\otimes\mathbf{M}_i)+\frac{1}{2}\left(\mathbf{I}\otimes\mathbf{P}_i^T\mathbf{P}_i\right)\right]\text{vec}\left(\bar{\mathbf{S}}^{(1)}-\bar{\mathbf{S}}^{(0)}\right)\leq 0. \tag{47}$$

where $\mathbf{e}_i = \mathrm{vec}(\mathbf{E}_i^T) \in \mathbb{C}^{n_T n_T \times 1}$, $\mathbf{r}_i = \mathrm{vec}(\mathbf{R}_i^T) \in \mathbb{C}^{n_T n_T \times 1}$, and κ_2 contains some constant terms that do not depend on $\{\mathbf{S}_i\}$. Let $\mathbf{s} = [\mathrm{vec}(\mathbf{S}_1)^T \mathrm{vec}(\mathbf{S}_2)^T \ldots \mathrm{vec}(\mathbf{S}_{|\mathcal{U}_I|})^T]^T \in \mathbb{C}^{n_T n_T |\mathcal{U}_I| \times 1}$. Note that $\mathrm{vec}(\bar{\mathbf{S}}) = \mathbf{T}\mathbf{s}$, where $\mathbf{T} \in \mathbb{C}^{n_T n_T \times n_T n_T |\mathcal{U}_I|}$ is composed of $|\mathcal{U}_I|$ identity matrices of size $n_T n_T \times n_T n_T$, i.e., $\mathbf{T} = [\mathbf{I}\ \mathbf{I}\ \ldots\ \mathbf{I}]$. Now, we can rewrite (55) as (omitting the constant terms)

$$\hat{R}_i(\mathbf{S}, \mathbf{S}^{(0)}) = \mathbf{s}^H \mathbf{T}^H \left(\mathbf{I} \otimes \mathbf{M}_i\right) \mathbf{T}\mathbf{s} + \mathbf{e}_i^T \mathbf{T}\mathbf{s} + \mathbf{r}_i^T \mathrm{vec}\left(\mathbf{S}_i\right). \tag{56}$$

We know proceed to formulate the objective function (denoted by $\bar{f}_0(\mathbf{S}, \mathbf{S}^{(0)})$ of problem (18) but substituting the bound that we just computed and considering the proximal term. If we incorporate all the terms (but omitting the constant ones) we have

$$\begin{aligned} \bar{f}_0 \left(\mathbf{S}, \mathbf{S}^{(0)}\right) = & \\ & \sum_{i \in \mathcal{U}_I} \omega_i \left(\mathbf{s}^H \mathbf{T}^H \left(\mathbf{I} \otimes \mathbf{M}_i\right) \mathbf{T}\mathbf{s} + \mathbf{e}_i^T \mathbf{T}\mathbf{s} + \mathbf{r}_i^T \mathrm{vec}\left(\mathbf{S}_i\right) \right) \\ & - \rho \left\| \mathbf{S}_i - \mathbf{S}_i^{(0)} \right\|_F^2 \end{aligned} \tag{57}$$

$$\begin{aligned} = & \ \mathbf{s}^H \mathbf{T}^H \tilde{\mathbf{M}} \mathbf{T}\mathbf{s} + \tilde{\mathbf{e}}^T \mathbf{T}\mathbf{s} + \hat{\mathbf{r}}^T \mathbf{s} - \rho \mathbf{s}^H \mathbf{s} + \rho \mathbf{s}^{(0),H} \mathbf{s} \\ & + \rho \mathbf{s}^H \mathbf{s}^{(0)} - \rho \mathbf{s}^{(0),H} \mathbf{s}^{(0)}, \end{aligned} \tag{58}$$

where $\tilde{\mathbf{M}} = \sum_{i \in \mathcal{U}_I} \omega_i (\mathbf{I} \otimes \mathbf{M}_i) \in \mathbb{C}^{n_T n_T \times n_T n_T}$, $\tilde{\mathbf{e}} = \sum_{i \in \mathcal{U}_I} \omega_i \mathbf{e}_i$, $\hat{\mathbf{r}} = [\mathbf{r}_1^T\ \mathbf{r}_2^T\ \ldots\ \mathbf{r}_{|\mathcal{U}_I|}^T]^T \in \mathbb{C}^{n_T n_T |\mathcal{U}_I| \times 1}$, and $\mathbf{s}^{(0)} = [\mathrm{vec}(\mathbf{S}_1^{(0)})^T \mathrm{vec}(\mathbf{S}_2^{(0)})^T \ldots \mathrm{vec}(\mathbf{S}_{|\mathcal{U}_I|}^{(0)})^T]^T \in \mathbb{C}^{n_T n_T |\mathcal{U}_I| \times 1}$. Now taking into account that the objective function $\bar{f}_0(\mathbf{S}, \mathbf{S}^{(0)})$ must be real and combining terms (omitting terms that do not depend on $\mathbf{s}$) we obtain

$$\bar{f}_0(\mathbf{S}, \mathbf{S}^{(0)}) = \mathbf{s}^H \mathbf{C}\mathbf{s} + \mathbf{b}^T \mathbf{s} + \mathbf{s}^H \mathbf{b}^*, \tag{59}$$

where $\mathbf{b}^T = \frac{1}{2}\tilde{\mathbf{e}}^T \mathbf{T} + \frac{1}{2}\hat{\mathbf{r}}^T + \rho \mathbf{s}^{(0),H} \in \mathbb{C}^{1 \times n_T n_T |\mathcal{U}_I|}$ and matrix $\mathbf{C}$ is $\mathbf{C} = \mathbf{T}^H \tilde{\mathbf{M}} \mathbf{T} - \rho \mathbf{I} \in \mathbb{C}^{n_T n_T |\mathcal{U}_I| \times n_T n_T |\mathcal{U}_I|}$. For convenient purposes, let us change the sign of $\bar{f}_0(\mathbf{S}, \mathbf{S}^{(0)})$ such that $\bar{\bar{f}}_0(\mathbf{S}, \mathbf{S}^{(0)}) = -\bar{f}_0(\mathbf{S}, \mathbf{S}^{(0)}) = \mathbf{s}^H \tilde{\mathbf{C}} \mathbf{s} - \mathbf{b}^T \mathbf{s} - \mathbf{s}^H \mathbf{b}^*$, where $\tilde{\mathbf{C}} = -\mathbf{C} \succeq 0$. Finally, we can equivalently rewrite the objective function as the following expression (with this new reformulation, the objective is to minimize $\bar{\bar{f}}_0(\mathbf{S}, \mathbf{S}^{(0)})$ instead of maximizing it):

$$\bar{\bar{f}}_0(\mathbf{S}, \mathbf{S}^{(0)}) = \|\tilde{\mathbf{C}}^{\frac{1}{2}} \mathbf{s} - \mathbf{c}\|_2^2, \tag{60}$$

where

$$\mathbf{c} = \tilde{\mathbf{C}}^{-\frac{1}{2}} \mathbf{b}^* \in \mathbb{C}^{n_T n_T |\mathcal{U}_I| \times 1}. \tag{61}$$

Note that the term $\mathbf{c}^H \mathbf{c}$ does not affect the optimum value of the optimization variables as this term does not depend on $\mathbf{s}$. Now, we can reformulate the optimization problem presented in (18) as

$$\begin{aligned} \underset{\{\mathbf{S}_i\}, \mathbf{s}}{\text{minimize}} \quad & \|\tilde{\mathbf{C}}^{\frac{1}{2}} \mathbf{s} - \mathbf{c}\|_2^2 \\ \text{subject to} \quad & C1: \mathbf{T}_i \mathbf{s} = \mathrm{vec}\left(\mathbf{S}_i\right), \qquad \forall i \in \mathcal{U}_I \\ & C2: \mathbf{S} \in \mathcal{S}_1, \end{aligned} \tag{62}$$

where $\mathbf{T}_i = [\underbrace{\mathbf{0}, \mathbf{0}, \ldots, \mathbf{0}}_{i-1}, \mathbf{I}, \mathbf{0}, \ldots, \mathbf{0}] \in \mathbb{R}^{n_T n_T \times n_T n_T |\mathcal{U}_I|}$ is composed of zero matrices of dimension $n_T n_T \times n_T n_T$ with an identity matrix at the i-th position. Problem (62) can be further reformulated as

$$\begin{aligned} \underset{\{\mathbf{S}_i\}, \mathbf{s}, t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & C1: \|\tilde{\mathbf{C}}^{\frac{1}{2}} \mathbf{s} - \mathbf{c}\|_2 \leq t \\ & C2: \mathbf{T}_i \mathbf{s} = \mathrm{vec}\left(\mathbf{S}_i\right), \qquad \forall i \in \mathcal{U}_I \\ & C3: \mathbf{S} \in \mathcal{S}_1, \end{aligned} \tag{63}$$

and, finally, as the following standard SDP optimization problem that can be solved fast with specific SDP solvers [32]:

$$\begin{aligned} \underset{\{\mathbf{S}_i\}, \mathbf{s}, t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & C1: \begin{bmatrix} t\mathbf{I} & \tilde{\mathbf{C}}^{\frac{1}{2}} \mathbf{s} - \mathbf{c} \\ \left(\tilde{\mathbf{C}}^{\frac{1}{2}} \mathbf{s} - \mathbf{c}\right)^H & 1 \end{bmatrix} \succeq 0 \\ & C2: \mathbf{T}_i \mathbf{s} = \mathrm{vec}\left(\mathbf{S}_i\right), \qquad \forall i \in \mathcal{U}_I \\ & C3: \mathbf{S} \in \mathcal{S}_1. \end{aligned} \tag{64}$$

Appendix D
Proof of Proposition 7

The proposed quadratic surrogate function of $s_i(\mathbf{S})$ has the following form:

$$\begin{aligned} \hat{s}_i(\mathbf{S}, \mathbf{S}^{(0)}) \triangleq & \log\det\left(\mathbf{I} + \mathbf{H}_i \sum_{k \in \mathcal{U}_I} \mathbf{S}_k^{(0)} \mathbf{H}_i^H\right) \\ & + \sum_{\ell \in \mathcal{U}_I} \mathrm{Re}\Big\{ \mathrm{Tr}\left(\mathbf{G}_{\ell i}\left(\mathbf{S}_\ell - \mathbf{S}_\ell^{(0)}\right)\right)\Big\} \\ & + \sum_{\ell \in \mathcal{U}_I} \mathrm{Tr}\left(\left(\mathbf{S}_\ell - \mathbf{S}_\ell^{(0)}\right)^H \mathbf{M}_{\ell i}\left(\mathbf{S}_\ell - \mathbf{S}_\ell^{(0)}\right)\right) \\ \leq & \ \log\det\left(\mathbf{I} + \mathbf{H}_i \sum_{k \in \mathcal{U}_I} \mathbf{S}_k \mathbf{H}_i^H\right), \quad \forall \mathbf{S}_\ell, \mathbf{S}_\ell^{(0)} \in \mathcal{S}_+^{n_T}, \end{aligned} \tag{65}$$

where matrices $\mathbf{G}_i \in \mathbb{C}^{n_T \times n_T}$ and $\mathbf{M}_i \in \mathbb{C}^{n_T \times n_T}$ need to be found such that conditions (A1) through (A4) are satisfied. Note that (A1) and (A4) are already satisfied. Only (A2) and (A3) must be ensured. Let us start with condition (A3). Let $\mathbf{S}_\ell^{(0)}$, $\mathbf{S}_\ell^{(1)} \in \mathcal{S}_+^{n_T}, \forall \ell$. Then, the directional derivative of the surrogate function $\hat{s}_i(\mathbf{S}, \mathbf{S}^{(0)})$ in (65) at $\mathbf{S}_\ell^{(0)}$ with direction $\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)}$ is given by

$$\sum_{\ell \in \mathcal{U}_I} \mathrm{Re}\Big\{ \mathrm{Tr}\left(\mathbf{G}_{\ell i}\left(\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)}\right)\right)\Big\}, \tag{66}$$

equation (67) as shown at the bottom of next page and the directional derivative of the right hand side of (65) at $\mathbf{S}_\ell^{(0)}$ with direction $\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)}$ is given by (67). From (66) and (67), we identify the matrices $\mathbf{G}_{\ell i}$ as

$$\mathbf{G}_{\ell i} = \mathbf{H}_i^H \left(\mathbf{I} + \mathbf{H}_i \sum_{k \in \mathcal{U}_I} \mathbf{S}_k^{(0)} \mathbf{H}_i^H \right)^{-1} \mathbf{H}_i, \quad \mathbf{G}_{\ell i} = \mathbf{G}_{\ell i}^H, \tag{68}$$

where we find that all matrices $\mathbf{G}_{\ell i}$ for a given user i can be the same, $\mathbf{G}_i = \mathbf{G}_{\ell i}$ (i.e., they do not depend on ℓ).

Now, we seek to find matrices $\{\mathbf{M}_{\ell i}\}$ based on condition (A2). To this end, we follow the same procedure presented before. We make linear cuts in each possible direction and apply the condition over the second derivative (see (42)). The second derivative of the left hand side of (65) is given by

$$2 \sum_{\ell \in \mathcal{U}_I} \mathrm{Tr} \left(\left(\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)} \right)^H \mathbf{M}_{\ell i} \left(\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)} \right) \right) = \tag{69}$$

$$2 \sum_{\ell \in \mathcal{U}_I} \mathrm{vec} \left(\left(\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)} \right)^T \right)^T (\mathbf{I} \otimes \mathbf{M}_{\ell i})\, \mathrm{vec} \left(\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)} \right),$$

equation (70) as shown at the bottom of this page and the second derivative of the right hand side is given by (70), where $\mathbf{P}_i = \mathbf{H}_i^H (\mathbf{I} + \mathbf{H}_i (\sum_{\ell \in \mathcal{U}_I} (\mathbf{S}_\ell^{(0)} + \mu (\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)}))) \mathbf{H}_i^H)^{-1} \mathbf{H}_i$, being constant $\mu \in [0, 1]$. Now, let $\mathbf{s} = [\mathrm{vec}(\mathbf{S}_1^{(1)} - \mathbf{S}_1^{(0)})^T \cdots \mathrm{vec}(\mathbf{S}_{|\mathcal{U}_I|}^{(1)} - \mathbf{S}_{|\mathcal{U}_I|}^{(0)})^T]^T$ and let us introduce the following block diagonal matrix

$$\tilde{\mathbf{M}}_i = \begin{bmatrix} \mathbf{I} \otimes \mathbf{M_{1i}} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \otimes \mathbf{M}_{2i} & & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{I} \otimes \mathbf{M}_{|\mathcal{U}_I| i} \end{bmatrix}. \tag{71}$$

Then we have that the following condition should be fulfilled:

$$2\mathbf{s}^H \tilde{\mathbf{M}}_i \mathbf{s} + \mathbf{s}^H \mathbf{T}^H \left(\mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i \right) \mathbf{T} \mathbf{s} \leq 0, \tag{72}$$

which means that

$$\tilde{\mathbf{M}}_i + \frac{1}{2} \mathbf{T}^H \left(\mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i \right) \mathbf{T} \preceq 0. \tag{73}$$

Note that the particular structure of matrix $\mathbf{T}^H \left(\mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i \right) \mathbf{T}$ is given by

$$\mathbf{T}^H \left(\mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i \right) \mathbf{T} = \begin{bmatrix} \mathbf{I} \otimes \mathbf{P_i^T P_i} & \dots & \mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i \\ \mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i & & \\ \vdots & \ddots & \vdots \\ \mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i & \dots & \mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i \end{bmatrix}, \tag{74}$$

From the previous conditions we can see that all matrices $\mathbf{M}_{\ell i}$ will be the same for user i, i.e., $\mathbf{M}_{\ell i} = \mathbf{M}_i$, $\forall \ell$. Now if we choose the particular structure $\mathbf{M}_i = \alpha_i \mathbf{I}$, then condition (73) is equivalent to

$$\alpha_i \mathbf{I} + \frac{1}{2} \mathbf{T}^H \left(\mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i \right) \mathbf{T} \preceq 0. \tag{75}$$

Now, condition (75) is equivalent to

$$\alpha_i \mathbf{g}^H \mathbf{g} \leq -\frac{1}{2} \mathbf{g}^H \mathbf{T}^H \left(\mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i \right) \mathbf{T} \mathbf{g}, \quad \forall \mathbf{g}. \tag{76}$$

If we propose a value of α such that

$$\alpha_i \mathbf{g}^H \mathbf{g} \leq -\frac{1}{2} \|\mathbf{T}\mathbf{g}\|_2^2 \lambda_{\max} \left(\mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i \right), \quad \forall \mathbf{g}, \tag{77}$$

$$\alpha_i \mathbf{g}^H \mathbf{g} \leq -\frac{1}{2} \|\mathbf{T}\mathbf{g}\|_2^2 \lambda_{\max} \left(\mathbf{P}_i^T \mathbf{P}_i \right), \quad \forall \mathbf{g}. \tag{78}$$

are fulfilled, this ensures that (76) is fulfilled. Therefore, the condition over α shown in (77) and (78) are sufficient conditions to fulfilled (75). Now, the term $\|\mathbf{T}\mathbf{g}\|_2^2$ can be further simplified. Based on the structure of matrix $\mathbf{T}$, we have that

$$\| \mathbf{T}\mathbf{g}\|_2^2 = \sum_{i=1}^{n_T n_T} |\mathbf{g}_i + \mathbf{g}_{i+n_T n_T + 1} + \ldots + \mathbf{g}_{i + n_T n_T (|\mathcal{U}_I| - 1) + 1}|^2 \tag{79}$$

$$\leq \sum_{i=1}^{n_T n_T} ||\mathcal{U}_I| \max\{\mathbf{g}_i, \ldots, \mathbf{g}_{i + n_T n_T (|\mathcal{U}_I| - 1) + 1}\}|^2 \tag{80}$$

$$\leq \sum_{i=1}^{n_T n_T} |\mathcal{U}_I|^2 \left(|\mathbf{g}_i|^2 + \ldots + |\mathbf{g}_{i + n_T n_T (|\mathcal{U}_I| - 1) + 1}|^2 \right) \tag{81}$$

$$= |\mathcal{U}_I|^2 \sum_{i=1}^{n_T n_T |\mathcal{U}_I|} |\mathbf{g}_i|^2 = |\mathcal{U}_I|^2 \|\mathbf{g}\|_2^2. \tag{82}$$

$$\mathrm{Tr} \left(\mathbf{H}_i^H \left(\mathbf{I} + \mathbf{H}_i \sum_{k \in \mathcal{U}_I} \mathbf{S}_k^{(0)} \mathbf{H}_i^H \right)^{-1} \mathbf{H}_i \left(\sum_{\ell \in \mathcal{U}_I} \left(\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)} \right) \right) \right)$$
$$= \sum_{\ell \in \mathcal{U}_I} \mathrm{Tr} \left(\mathbf{H}_i^H \left(\mathbf{I} + \mathbf{H}_i \sum_{k \in \mathcal{U}_I} \mathbf{S}_k^{(0)} \mathbf{H}_i^H \right)^{-1} \mathbf{H}_i \left(\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)} \right) \right) \tag{67}$$

$$\mathrm{vec} \left(\left(\sum_{\ell \in \mathcal{U}_I} \left(\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)} \right) \right)^T \right)^T \left(\mathbf{I} \otimes \mathbf{P}_i^T \mathbf{P}_i \right) \mathrm{vec} \left(\sum_{\ell \in \mathcal{U}_I} \left(\mathbf{S}_\ell^{(1)} - \mathbf{S}_\ell^{(0)} \right) \right), \tag{70}$$

Thus, a sufficient condition to fulfill (78) is

$$\alpha_i \|\mathbf{g}\|_2^2 \leq -\frac{1}{2}|\mathcal{U}_I|^2 \|\mathbf{g}\|_2^2 \lambda_{\max}\left(\mathbf{P}_i^T \mathbf{P}_i\right), \quad \forall \mathbf{g}, \tag{83}$$

and, finally,

$$\alpha_i \leq -\frac{1}{2}|\mathcal{U}_I|^2 \lambda_{\max}\left(\mathbf{P}_i^T \mathbf{P}_i\right) \leq -\frac{1}{2}|\mathcal{U}_I|^2 \lambda_{\max}^2(\mathbf{H}_i^H \mathbf{H}_i). \tag{84}$$

Hence, a possible matrix $\mathbf{M}_i$ satisfying assumptions $(\mathrm{A1})-(\mathrm{A4})$ is, finally,

$$\mathbf{M}_i = -\frac{1}{2}|\mathcal{U}_I|^2 \lambda_{\max}^2(\mathbf{H}_i^H \mathbf{H}_i)\mathbf{I}. \tag{85}$$

References

[1] X. Lu *et al.*, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 757–789, Apr.–Jun. 2015.

[2] J. Paradiso and T. Starner, "Energy scavenging for mobile wireless electronics," *IEEE Comput. Pervasive*, vol. 4, no. 1, pp. 18–27, Jan. 2005.

[3] S. Bi, C. K. Ho, and R. Zhang, "Wireless powered communication: Opportunities and challenges," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 117–125, Apr. 2015.

[4] L. R. Varshney, "Transporting information and energy simultaneously," in *Proc. Int. Symp. Inf. Theory*, Toronto, ON, Canada, Jul. 2008, pp. 1612–1616.

[5] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 1989–2001, May 2013.

[6] J. Rubio and A. Pascual-Iserte, "Simultaneous wireless information and power transfer in multiuser MIMO systems," in *Proc. IEEE Global Commun. Conf.*, Atlanta, GA, USA, Dec. 2013, pp. 2755–2760.

[7] Q. H. Spencer *et al.*, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, Feb. 2004.

[8] J. Park and B. Clerckx, "Joint wireless information and energy transfer in a two-user MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 12, no. 8, pp. 4210–4221, Aug. 2013.

[9] J. Park and B. Clerckx, "Joint wireless information and energy transfer in a K-user MIMO interference channel," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5781–5796, Oct. 2014.

[10] Z. Zong *et al.*, "Optimal transceiver design for SWIPT in K-user MIMO interference channels," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 430–445, Jan. 2016.

[11] J. Xu, L. Liu, and R. Zhang, "Multiuser MISO beamforming for simultaneous wireless information and power transfer," *IEEE Trans. Signal Process.*, vol. 62, no. 18, pp. 4798–4810, Sep. 2015.

[12] Q. Shi, L. Liu, W. Xu, and R. Zhang, "Joint transmit beamforming and receive power splitting for MISO SWIPT systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 6, pp. 3269–3280, Jun. 2014.

[13] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.

[14] X. Gui, Z. Zhu, and I. Lee, "Sum rate maximizing in a multi-user MIMO system with SWIPT," in *Proc. IEEE Veh. Technol. Conf.*, Singapore, May 2015, pp. 1–5.

[15] S. Christensen, R. Agarwal, E. Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[16] J. Rubio, A. Pascual-Iserte, D. P. Palomar, and A. Goldsmith, "SWIPT techniques for multiuser MIMO broadcast systems," in *Proc. IEEE Int. Symp. Pers. Indoor Mobile Radio Commun.*, Sep. 2016.

[17] M. Ehrgott, *Multicriteria Optimization*. New York, NY, USA: Springer-Verlag, 2005.

[18] I. Kim and O. de Weck, "Adaptive weighted sum method for multiobjective optimization: a new method for Pareto front generation," *Struct. Multidiscip. Optim.*, vol. 31, no. 2, pp. 105–116, Feb. 2006.

[19] D. R. Hunter and K. Lange, "A tutorial on MM algorithms," *Amer. Statistician*, vol. 58, no. 1, pp. 30–37, Feb. 2004.

[20] Z. Ding *et al.*, "Application of smart antenna technologies in simultaneous wireless information and power transfer," *IEEE Commun. Mag.*, vol. 53, no. 4, pp. 86–93, Apr. 2015.

[21] E. Boshkovska, D. Ng, N. Zlatanov, and R. Schober, "Practical non-linear energy harvesting model and resource allocation for SWIPT systems," *IEEE Commun. Lett.*, vol. 19, no. 12, pp. 2082–2085, Dec. 2015.

[22] L. Liu, R. Zhang, and K.-C. Chua, "Wireless information and power transfer: A dynamic power splitting approach," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3990–4001, Sep. 2013.

[23] L. Liu, Y.-H. Nam, and J. Zhang, "Proportional fair scheduling for multi-cell multi-user MIMO systems," in *Proc. 44th Annu. Conf. Inf. Sci. Syst.*, Princeton, NJ, USA, Mar. 2010, pp. 1–6.

[24] M. Andrews *et al.*, "Providing quality of service over a shared wireless link," *IEEE Commun. Mag.*, vol. 39, no. 2, pp. 150–154, Feb. 2001.

[25] M. Hong, Q. Li, and Y.-F. Liu, "Decomposition by successive convex approximation: A unifying approach for linear transceiver design in heterogeneous networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1377–1392, Feb. 2016.

[26] G. Scutari, F. Facchinei, P. Song, D. P. Palomar, and J.-S. Pang, "Decomposition by partial linearization: Parallel optimization of multi-agent systems," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 641–656, Feb. 2014.

[27] S. You, L. Chen, and Y. E. Liu, "Convex-concave procedure for weighted sum-rate maximization in a MIMO interference network," in *Proc. IEEE Glob. Commun. Conf.*, Austin, TX, USA, Dec. 2014, pp. 4060–4065.

[28] D. Henrion and J. Malick, "Projection methods in conic optimization," *Int. Series Oper. Res. Manage. Sci.*, vol. 166, pp. 565–600, Sep. 2011.

[29] S. Boyd and A. Mutapcic, "Subgradient methods," Apr. 2008. [Online]. Available: https://see.stanford.edu/materials/lsocoee364b/02-subgrad_method_notes.pdf

[30] Q. Shi, W. Xu, J. Wu, E. Song, and Y. Wang, "Secure beamforming for MIMO broadcasting with wireless information and power transfer," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2841–2853, May 2015.

[31] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming," Vers. 2.0 beta, Sep. 2013. [Online]. Available: http://cvxr.com/cvx

[32] J. F. Sturm, "Sedumi software." [Online]. Available: http://sedumi.ie.lehigh.edu

[33] 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.

[34] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Nov. 2010.

[35] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus With Application in Statistics and Econometrics*. Essex, U.K.: Wiley, 1988.

[36] B. Wang and F. Zhang, "Some inequalities for the eigenvalues of the product of positive semidefinite Hermitian matrices," *Struct. Multidiscip. Optim.*, vol. 160, pp. 113–118, Jan. 1992.

Javier Rubio received the B.S. (with highest hons.), M.S. (with highest honors), and Ph.D. degrees (cum laude), all in electrical engineering, from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in July 2010, July 2012, June 2016, respectively. From June 2015 to September 2015 he was with the Wireless Systems Lab at Stanford University under the supervision of Prof. Andrea Goldsmith. From September 2009 to July 2010, he was with the Wireless Access Research Center, University of Limerick, Limerick, Ireland, where he developed his bachelor thesis in the field of cognitive radio networks. In April 2011, he joined the Department of Signal Theory and Communications at UPC where he worked as a Research Assistant until December 2012. His main research interests include energy-aware resource allocation, energy harvesting techniques, optimization theory, and heterogeneous networks.

Antonio Pascual-Iserte (S'01–M'07–SM'11) was born in Barcelona, Spain, in 1977. He received the Electrical Engineering and Ph.D. degrees from the Universitat Politècnica de Catalunya (UPC), Barcelona, Spain, in September 2000 and February 2005, respectively. From September 1998 to June 1999, he worked as a Teaching Assistant in the field of microprocessor programming in the Electronic Engineering Department, UPC, and from June 1999 to December 2000 he was with Retevision R&D, working on the implantation of the DVB-T and T-DAB networks in Spain. In January 2001, he joined the Department of Signal Theory and Communications, UPC, where he worked as a Research Assistant until September 2003. He received a predoctoral grant from the Catalan government for the Ph.D. studies during this period. He became an Assistant Professor in September 2003 and since April 2008 he is an Associate Professor. He currently teaches undergraduate courses in linear systems and signal theory. He also teaches post-graduate courses in advanced signal processing and estimation theory in the Department of Signal Theory and Communications. His current research interests include array processing, robust designs, OFDM, MIMO channels, multiuser access, 5G, stochastic geometry, HetNets, and optimization theory. He has been involved in several research projects funded by the Spanish Government and the European Commission. He has also published several papers in international and national conference and journals. He was awarded with the First National Prize of 2000/2001 University Education by the Spanish Ministry of Education and Culture, and with the Best 2004/2005 Ph.D. Thesis Prize by UPC.

Daniel P. Palomar (S'99–M'03–SM'08–F'12) received the Electrical Engineering and Ph.D. degrees from the Technical University of Catalonia (UPC), Barcelona, Spain, in 1998 and 2003, respectively. He is a Professor in the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology (HKUST), Clear Water Bay, Hong Kong, where he joined in 2006. He had previously held several research appointments, namely, at King's College London, London, U.K.; Stanford University, Stanford, CA, USA; Telecommunications Technological Center of Catalonia, Barcelona; Royal Institute of Technology (KTH), Stockholm, Sweden; University of Rome La Sapienza, Rome, Italy; and Princeton University, Princeton, NJ, USA. His current research interests include applications of convex optimization theory, game theory, and variational inequality theory to financial systems, big data systems, and communication systems. Since 2013, he has been a Fellow of the Institute for Advance Study, HKUST. He received the 2004/06 Fulbright Research Fellowship, the 2004 and 2015 (co-author) Young Author Best Paper Awards by the IEEE Signal Processing Society, the 2015–2016 HKUST Excellence Research Award, the 2002/03 best Ph.D. prize in Information Technologies and Communications by the UPC, the 2002/03 Rosina Ribalta first prize for the Best Doctoral Thesis in Information Technologies and Communications by the Epson Foundation, and the 2004 prize for the best Doctoral Thesis in Advanced Mobile Communications by the Vodafone Foundation and COIT. He is a Guest Editor of the IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING 2016 Special Issue on Financial Signal Processing and Machine Learning for Electronic Trading and has been an Associate Editor of IEEE TRANSACTIONS ON INFORMATION THEORY and of IEEE TRANSACTIONS ON SIGNAL PROCESSING, a Guest Editor of the IEEE SIGNAL PROCESSING MAGAZINE 2010 Special Issue on Convex Optimization for Signal Processing, the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2008 Special Issue on Game Theory in Communication Systems, and the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS 2007 Special Issue on Optimization of MIMO Transceivers for Realistic Communication Networks.

Andrea Goldsmith (S'90–M'93–SM'99–F'05) received the B.S., M.S., and Ph.D. degrees, all in electrical engineering, from University of California, Berkeley, CA, USA. She is the Stephen Harris Professor in the School of Engineering and a Professor of electrical engineering at Stanford University. She was previously on the Faculty of Electrical Engineering at Caltech. She co-founded and served as a Chief Scientist of Plume WiFi, and also co-founded and served as the CTO of Quantenna Communications, Inc. She has held industry positions at Maxim Technologies, Memorylink Corporation, and AT&T Bell Laboratories. She is author of the book *Wireless Communications* (Cambridge University Press) and co-author of the books *MIMO Wireless Communications* (Cambridge University Press) and *Principles of Cognitive Radio* (Cambridge University Press), as well as an inventor on 28 patents. Her research interests include information theory and communication theory, and their application to wireless communications and related fields. She has served on the Steering Committee for the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONSand as Editor for the IEEE TRANSACTIONS ON INFORMATION THEORY, the *Journal on Foundations and Trends in Communications and Information Theory and in Networks*, the IEEE TRANSACTIONS ON COMMUNICATIONS, and the IEEE WIRELESS COMMUNICATIONS MAGAZINE. She participates actively in committees and conference organization for the IEEE Information Theory and Communications Societies and has served on the Board of Governors for both societies. She has also been a Distinguished Lecturer for both societies, served as the President of the IEEE Information Theory Society in 2009, founded and chaired the student committee of the IEEE Information Theory society, and chaired the Emerging Technology Committee of the IEEE Communications Society. She currently chairs the IEEE ad hoc committee on women and URMs, and the Women in Technology Leadership Roundtable working group on metrics. She served as the Chair of Stanfords Faculty Senate and for multiple terms as a Senator, and currently serves on its Budget Group, Committee on Research, and Task Force on Women and Leadership. She is a Fellow of Stanford, and has received several awards for her work, including the inaugural University Postdoc Mentoring Award, the IEEE ComSoc Edwin H. Armstrong Achievement Award as well as Technical Achievement Awards in Communications Theory and in Wireless Communications, the National Academy of Engineering Gilbreth Lecture Award, the IEEE ComSoc and Information Theory Society Joint Paper Award, the IEEE ComSoc Best Tutorial Paper Award, the Alfred P. Sloan Fellowship, the WICE Technical Achievement Award, and the Silicon Valley/San Jose Business Journals Women of Influence Award.