SWIPT Techniques for Multiuser MIMO Broadcast Systems

Javier Rubio⋆, Antonio Pascual-Iserte⋆,
Daniel P. Palomar†, and Andrea Goldsmith‡
⋆Universitat Politècnica de Catalunya (UPC), Barcelona, Spain
†Hong Kong University of Science and Technology (HKUST), Hong Kong
‡Stanford University, CA, USA
Emails:{javier.rubio.lopez, antonio.pascual}@upc.edu, palomar@ust.hk, andrea@wsl.stanford.edu

***Abstract*—In this paper, 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 design is formulated as a nonconvex optimization problem in which system sum rate is optimized considering per-user harvesting constraints. Two different approaches are proposed. The first approach is based on a classical gradient-based method for constrained optimization. The second approach is based on difference of convex (DC) programming. The idea behind this approach is to obtain a convex function that approximates the nonconvex objective and, then, solve a series of convex subproblems that, eventually, will provide a (locally) optimum solution of the general nonconvex problem. The solution obtained from the proposed approach is compared to the classical block-diagonalization (BD) strategy, typically used to solve the nonconvex multiuser MIMO network by forcing no inter-user interference. Simulation results show that the proposed approach improves both the system sum rate and the power harvested by users simultaneously. In terms of computational time, the proposed DC programming outperforms the classical gradient methods.**

I. Introduction

Simultaneous wireless information and power transfer (SWIPT) is a technique by which a transmitter actively feeds a receiver (or a set of receivers) with power that is sent through radio frequency (RF) signals and, simultaneously, sends useful information to the same or different receivers [1]. In this context, battery-constrained devices are able to prolong their operation time by means of recharging their batteries thanks to this energy harvesting process [2].

The concept of SWIPT was first studied from a theoretical point of view by Varshney [3]. He showed that, for the single-antenna additive white Gaussian noise (AWGN) channel, there exists a nontrivial trade-off in maximizing the data rate versus the power transmission. In [4], authors considered a single user multiple-input multiple-output (MIMO) scenario with one transmitter capable of transmitting information and power simultaneously to one receiver. Later, in [5], authors extended the work in [4] by considering that multiple users were present in the broadcast MIMO system. However, the multi-stream transmit covariance optimization that appears in broadcast MIMO systems is a very difficult nonconvex optimization problem. In order to overcome that difficulty, authors in [5] assumed a block-diagonalization (BD) strategy [6], in which interference among users is pre-canceled at the transmitter. The BD technique allows for a simple solution but wastes some degrees of freedom and, thus, degrades the overall performance. Paper [7], considered a MIMO network consisting of k transmitter-receiver pairs with co-channel interference. In [8], authors considered a MIMO system with single-stream transmission. In contrast to previous works where the system rate was optimized, their objective was to minimize the overall power consumption with signal to interference and noise ratio (SINR) constraints and per-user harvesting constraints. Multiuser broadcast networks can also be found under the framework of multiple-input single-output (MISO) beamforming as in [9]. The main difference of our work with respect to the previous works is that we assume a broadcast multiuser multi-stream (not BD-based) MIMO SWIPT network, which is a scenario not considered before.

The scope of this paper is to generalize all the previous works (specially [4] and [5]). The approach followed in this paper is the same as the one by the same authors in [10]. We consider a multiuser multi-stream MIMO SWIPT network. We assume that interference is not pre-canceled (that is, BD 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 gradient techniques and on difference of convex (DC) programming [11]. The gradient techniques developed in this paper are then used in the journal version [10] as benchmarks. Additionally, in the journal version we extend the DC method presented in this paper and consider other problem formulations.

The remainder of this paper is organized as follows. In Section II, we introduce the system model. Section III is devoted to presenting the problem formulation. In Section IV, we derive gradient-based techniques to solve the nonconvex SWIPT problem. In Section V, we develop an approach based on DC

The research leading to these results has received funding from the Spanish Ministry of Economy and Competitiveness (Ministerio de Economía y Competitividad) through the project TEC2011-29006-C03-02 (GRE3N-LINK-MAC), project TEC2013-41315-R (DISNET), and FPI grant BES-2012-052850, from the Catalan Government (AGAUR) through the grant 2014 SGR 60, and from the Hong Kong Government through the research grant Hong Kong RGC 16207814.

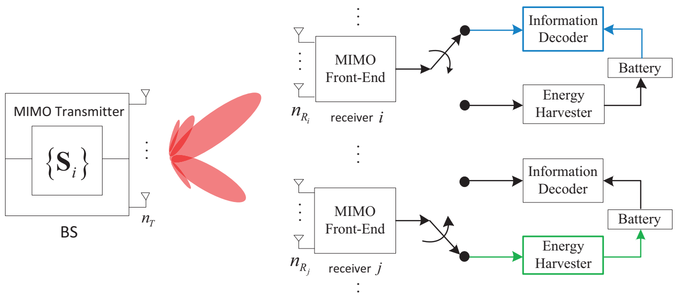

Fig. 1: Schematic representation of the downlink broadcast multiuser communication system. Note that each user can switch from an information decoder receiver to an energy harvester receiver.

programming to tackle the SWIPT problem. Section VI evaluates numerically the performance of the previous approaches and, finally, Section VII presents some conclusions.

II. System Model

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. The BS transmits information signals to some receivers whereas other receivers use those signals to recharge their batteries through energy harvesting. We assume that a given user is not able to decode information and to harvest energy simultaneously. 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 $|\mathcal{U}_I| + |\mathcal{U}_E| = N + M = K$.[1] 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 proposed system is depicted in Fig. 1.

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 out of the scope of the paper).

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

$$\mathbf{y}_i = \mathbf{H}_i \mathbf{B}_i \mathbf{x}_i + \mathbf{H}_i \sum_{\substack{k \in \mathcal{U}_I \\ k \neq i}} \mathbf{B}_k \mathbf{x}_k + \mathbf{n}_i, \quad \forall i \in \mathcal{U}_I. \tag{1}$$

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 considered Gaussian with $\mathbb{E}\left[\mathbf{n}_i \mathbf{n}_i^H\right] = \mathbf{I}_{n_{R_i}}$[2]. Note that the middle term of (1) is an interference term. The covariance matrix of the interference plus noise is written as

$$\boldsymbol{\Omega}_i(\mathbf{S}_{-i}) = \mathbf{H}_i \mathbf{S}_{-i} \mathbf{H}_i^H + \mathbf{I}, \quad \forall i \in \mathcal{U}_I, \tag{2}$$

where $\mathbf{S}_{-i} = \sum_{k \in \mathcal{U}_I, k \neq i} \mathbf{S}_k$.

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, i.e.,

$$\begin{aligned} \bar{Q}_j &= \zeta_j \mathbb{E}\left[\left\| \mathbf{H}_j \sum_{i \in \mathcal{U}_I} \mathbf{B}_i \mathbf{x}_i \right\|^2\right] \\ &= \zeta_j \sum_{i \in \mathcal{U}_I} \mathbb{E}[\|\mathbf{H}_j \mathbf{B}_i \mathbf{x}_i\|^2], \quad \forall j \in \mathcal{U}_E, \end{aligned} \tag{3}$$

where ζ_j is a constant that accounts for the loss in the transducer for converting the harvested RF power to electrical power to charge the battery. Notice that, for simplicity, in (3) we have omitted the harvested power due to the noise term since it can be assumed negligible.

Let $\tilde{\mathbf{x}} = \mathbf{B}\mathbf{x}$ denote the signal vector transmitted by the BS, where the joint precoding matrix is defined as $\mathbf{B} = [\mathbf{B}_1 \ \ldots \ \mathbf{B}_N] \in \mathbb{C}^{n_T \times n_S}$, being $n_S = \sum_{i \in \mathcal{U}_I} n_{S_i}$ the total number of streams of all information users, and the data vector as $\mathbf{x} = \left[\mathbf{x}_1^T \ \ldots \ \mathbf{x}_N^T\right]^T \in \mathbb{C}^{n_S \times 1}$, 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.

III. Problem Formulation

In this section, we consider the design of the covariance matrices $\{\mathbf{S}_i\}$ based on the maximization of the sum-rate with individual power harvesting constraints. The optimization problem can be written as

$$\begin{aligned} \underset{\{\mathbf{S}_i\}}{\text{maximize}} \quad & \sum_{i \in \mathcal{U}_I} \omega_i R_i(\mathbf{S}) \\ \text{subject to} \quad & C1: \sum_{i \in \mathcal{U}_I} \mathrm{Tr}(\mathbf{H}_j \mathbf{S}_i \mathbf{H}_j^H) \geq Q_j, \quad \forall j \in \mathcal{U}_E \\ & C2: \sum_{i \in \mathcal{U}_I} \mathrm{Tr}(\mathbf{S}_i) \leq P_T \\ & C3: \mathbf{S}_i \succeq 0, \qquad \forall i \in \mathcal{U}_I, \end{aligned} \tag{4}$$

where $Q_j = \frac{\bar{Q}_j^{\min}}{\zeta_j}$, being $\{\bar{Q}_j^{\min}\}$ the set of minimum power harvesting constraints, ω_i are some real non-negative weights, $\mathbf{S} \triangleq (\mathbf{S}_i)_{\forall i \in \mathcal{U}_I}$, and the data rate expression, $R_i(\mathbf{S})$, is given by

$$\begin{aligned} R_i(\mathbf{S}) &= \log\det\left(\mathbf{I} + \mathbf{H}_i \mathbf{S}_i \mathbf{H}_i^H \boldsymbol{\Omega}_i^{-1}(\mathbf{S}_{-i})\right) \quad (5) \\ &= \underbrace{\log\det\left(\mathbf{I} + \mathbf{H}_i \bar{\mathbf{S}} \mathbf{H}_i^H\right)}_{\triangleq \, s_i(\mathbf{S})} - \underbrace{\log\det\left(\boldsymbol{\Omega}_i(\mathbf{S}_{-i})\right)}_{\triangleq \, g_i(\boldsymbol{\Omega}_i(\mathbf{S}_{-i}))}, \quad (6) \end{aligned}$$

being $\bar{\mathbf{S}} = \sum_{k \in \mathcal{U}_I} \mathbf{S}_k$.

The previous optimization problem in (4) is not convex due the objective functions (in particular, due to $\boldsymbol{\Omega}_i(\mathbf{S}_{-i})$) and is difficult to solve. Notice, in fact, that in the literature the interference in (2) is generally assumed not to exist as the transmission strategy cancels it out, thus, making the problem convex and easier to solve [6]. If the problem is not convex, the KKT conditions are generally neither necessary nor sufficient, so a local point cannot be obtained directly by solving them [12]. In this sense,

[1] 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.

[2] 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.

we propose two different methods to find local optima of (4). The first method is based on applying a gradient-type approach for constrained optimization problems. The second approach is based on DC programming [11]. Note that function $s_i(\mathbf{S})$ is concave with respect to (w.r.t.) $\mathbf{S}$ whereas function $-g_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}))$ is convex w.r.t. $\mathbf{S}$. Hence, function $R_i(\mathbf{S})$ is categorized as a DC function. In this case, what we propose is to concavify the non-concave function $-g_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}))$ following the approach in [13], in which the non-concave function is approximated by a linear (and, thus, concave) function.

IV. Gradients-Based Methods to Solve (4)

A. Gradient Method for Constrained Optimization

Let us first consider a gradient method applied directly to the problem in (4). The update equation based on the gradient for constrained optimization problem reads as

$$\mathbf{S}_i^{(q+1)} = \mathbf{S}_i^{(q)} + \alpha^{(q)}\mathbf{T}_i^{(q)}, \quad \forall i \in \mathcal{U}_I, \tag{7}$$

where

$$\mathbf{T}_i^{(q)} = \begin{cases} \nabla f_0\left(\mathbf{S}^{(q)}\right), & \text{If } \mathbf{S}^{(q)} \text{ is feasible}, \\ -\nabla f_j\left(\mathbf{S}^{(q)}\right), & \text{otherwise}, \end{cases} \tag{8}$$

where $\nabla f_0\left(\mathbf{S}^{(q)}\right)$ denotes the gradient of $f_0(\cdot)$ at point $\mathbf{S}^{(q)}$ (being $\mathbf{S}^{(q)} \triangleq (\mathbf{S}_i^{(q)})_{\forall i\in\mathcal{U}_I}$), $f_0(\cdot)$ is the objective function of problem (4), i.e., $f_0(\mathbf{S}) = \sum_{i\in\mathcal{U}_I} \omega_i s_i(\mathbf{S}) - \omega_i g_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}))$, and $f_j(\cdot)$ corresponds to any violated constraint in problem (13), and $\alpha^{(q)}$ is the step size chosen such that the diminishing conditions are fulfilled, i.e., $\lim_{q\to\infty} \alpha^{(q)} = 0,\ \sum_{q=1}^{\infty} \alpha^{(q)} = \infty$ [12]. In other words, if the current point is feasible, we use a subgradient based on the objective function and if the current point is infeasible, we choose any violated constraint and use a subgradient of the associated constraint function. In the latter case, we can choose any of the violated constraints, if there is more than one. The gradients are calculated in the Appendix. The overall algorithm is presented in Algorithm 1.

Algorithm 1 Algorithm for Solving Problem (4)

1: Repeat
2: Update: $\mathbf{S}_i^{(q+1)} = \mathbf{S}_i^{(q)} + \alpha^{(q)}\mathbf{T}_i^{(q)}$. Set $q = q + 1$
3: Until convergence of $\mathbf{S}_i^{(q)}$

One important remark to note is that the previous approach does not need to provide feasible points $\mathbf{S}_i^{(k,q)}$ at each iteration q, that is, the procedure may give points that violate the constraints at some intermediate iterations. That is not a problem since the optimal point $\mathbf{S}_i^{\star}$ provided once the algorithm has converged will satisfy all the constraints [12].

B. Projected Gradient Method for Constrained Optimization

Another approach is to apply a gradient method but assuring that the new generated iterates $\mathbf{S}_i^{(q+1)}$ fulfill the constraints at each iteration q. To do this, we propose to use a projected gradient method, which is given by

$$\mathbf{S}_i^{(q+1)} = \Pi\left(\mathbf{S}_i^{(q)} + \alpha_i^{(q)}\mathbf{Z}_i^{(q)}\right), \quad \forall i \in \mathcal{U}_I, \tag{9}$$

where Π is the projection on $\mathcal{S}$ and $\mathbf{Z}_i^{(q)} = \nabla f_0\left(\mathbf{S}^{(q)}\right)$. Let $\widehat{\mathbf{S}}_i^{(q)} = \mathbf{S}_i^{(q)} + \alpha^{(q)}\mathbf{Z}_i^{(q)}, \quad \forall i \in \mathcal{U}_I$. Now, to obtain the projector function we must solve the following optimization problem:

$$\begin{aligned} \underset{\left\{\mathbf{S}_i^{(q+1)}\right\}}{\text{minimize}} \quad & \max_{i\in\mathcal{U}_I} \left\|\mathbf{S}_i^{(q+1)} - \widehat{\mathbf{S}}_i^{(q)}\right\|_2^2 \\ \text{subject to} \quad & C1: \sum_{i\in\mathcal{U}_I} \mathrm{Tr}(\mathbf{H}_j\mathbf{S}_i\mathbf{H}_j^H) \geq Q_j, \quad \forall j \in \mathcal{U}_E \\ & C2: \sum_{i\in\mathcal{U}_I} \mathrm{Tr}(\mathbf{S}_i) \leq P_T \\ & C3: \mathbf{S}_i \succeq 0, \qquad \forall i \in \mathcal{U}_I, \end{aligned} \tag{10}$$

which can be reformulated as the following easy-to-solve semidefinite programming (SDP):

$$\begin{aligned} \underset{t,\left\{\mathbf{S}_i^{(q+1)}\right\}}{\text{minimize}} \quad & t \\ \text{subject to} \quad & C1: \begin{bmatrix} t\mathbf{I} & \mathbf{S}_i^{(q+1)} - \widehat{\mathbf{S}}_i^{(q)} \\ \mathbf{S}_i^{(q+1)} - \widehat{\mathbf{S}}_i^{(q)} & t\mathbf{I} \end{bmatrix} \succeq 0 \\ & C2: \sum_{i\in\mathcal{U}_I} \mathrm{Tr}(\mathbf{H}_j\mathbf{S}_i\mathbf{H}_j^H) \geq Q_j, \qquad \forall j \in \mathcal{U}_E \\ & C3: \sum_{i\in\mathcal{U}_I} \mathrm{Tr}(\mathbf{S}_i) \leq P_T \\ & C4: \mathbf{S}_i \succeq 0, \qquad \forall i \in \mathcal{U}_I. \end{aligned} \tag{11}$$

The overall algorithm is presented in Algorithm 2.

Algorithm 2 Algorithm for Solving Problem (4)

1: Repeat
2: Update: $\widehat{\mathbf{S}}_i^{(q)} = \mathbf{S}_i^{(q)} + \alpha^{(q)}\mathbf{Z}_i^{(q)}$
3: Solve optimization problem (11) $\longrightarrow \mathbf{S}_i^{(q+1)}$
Set $q = q + 1$
4: Until convergence of $\mathbf{S}_i^{(q)}$

V. DC Programming Method to Solve (4)

Motivated by the work in [13], in this approach, we derive a (linear) approximation for the non-concave function $-g_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}))$, in such a way that the modified problem is convex[3]. To this end, we derive a simple local approximation of $f_0(\mathbf{S}) = \sum_{i\in\mathcal{U}_I} \omega_i s_i(\mathbf{S}) - \omega_i g_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}))$. In order to find a concave lower bound of $f_0(\mathbf{S})$, $g_i(\cdot)$ can be upper bounded linearly at point $\mathbf{\Omega}_i^{(0)} = \sum_{k\in\mathcal{U}_I, k\neq i} \mathbf{H}_i\mathbf{S}_k^{(0)}\mathbf{H}_i^H + \mathbf{I}$ as

$$\begin{aligned} g_i(\mathbf{\Omega}_i(\mathbf{S}_{-i})) &\leq g_i\left(\mathbf{\Omega}_i^{(0)}\right) + \mathrm{Tr}\left(\left(\mathbf{\Omega}_i^{(0)}\right)^{-1}\left(\mathbf{\Omega}_i(\mathbf{S}_{-i}) - \mathbf{\Omega}_i^{(0)}\right)\right) \\ &= \text{constant} + \mathrm{Tr}\left(\left(\mathbf{\Omega}_i^{(0)}\right)^{-1}\mathbf{\Omega}_i(\mathbf{S}_{-i})\right) \\ &\triangleq \hat{g}_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}), \mathbf{\Omega}_i^{(0)}). \end{aligned} \tag{12}$$

Note that the upper bound $\hat{g}_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}), \mathbf{\Omega}_i^{(0)})$ can be used to build a lower bound of $f_0(\{\mathbf{S}_i\})$. By applying a successive approximation of $f_0(\cdot)$ through the application of the function $\hat{f}_0(\mathbf{S}, \mathbf{S}^{(k)}) =$

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

$\sum_{i\in\mathcal{U}_I} \omega_i s_i(\mathbf{S}) - \omega_i \hat{g}_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}), \mathbf{\Omega}_i^{(k)}) - \rho \left\| \mathbf{S}_i - \mathbf{S}_i^{(k)} \right\|_F^2$ (where $\mathbf{S}^{(k)} \triangleq (\mathbf{S}_i^{(k)})_{\forall i\in\mathcal{U}_I}$), we obtain an iterative algorithm based on the approach presented in [13] that converges to a stationary point (or local optimum) of the original problem (4). Note that 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. Given this, the optimization problem to be solved is

$$\begin{aligned} \underset{\{\mathbf{S}_i\}}{\text{maximize}} \; & \sum_{i\in\mathcal{U}_I} \omega_i s_i(\mathbf{S}) - \omega_i \hat{g}_i(\mathbf{\Omega}_i(\mathbf{S}_{-i}), \mathbf{\Omega}_i^{(k)}) - \rho \left\| \mathbf{S}_i - \mathbf{S}_i^{(k)} \right\|_F^2 \\ \text{subject to} \;\; & C1: \sum_{i\in\mathcal{U}_I} \mathrm{Tr}(\mathbf{H}_j \mathbf{S}_i \mathbf{H}_j^H) \geq Q_j, \quad \forall j \in \mathcal{U}_E \qquad (13) \\ & C2: \sum_{i\in\mathcal{U}_I} \mathrm{Tr}(\mathbf{S}_i) \leq P_T \\ & C3: \mathbf{S}_i \succeq 0, \qquad \forall i \in \mathcal{U}_I. \end{aligned}$$

The previous optimization problem is convex and can be solved using any standard convex optimization tools [12]. In order to obtain a (local) solution of (4), we must proceed iteratively until convergence of $\{\mathbf{S}_i^{(k)}\}$ is reached. The procedure is presented in Algorithm 3.

Algorithm 3 Algorithm for Solving Problem (4)

Initialize $\mathbf{S}^{(0)}$. Set $k = 0$
Repeat
  Generate $(k+1)$-th tuple $(\mathbf{S}_i^{\star})_{\forall i\in\mathcal{U}_I}$ by solving (13)
  Set $\mathbf{S}_i^{(k+1)} = \mathbf{S}_i^{\star}$, $\forall i \in \mathcal{U}_I$, and set $k = k + 1$
Until convergence is reached

VI. 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 this simulation, we consider a system composed of 1 transmitter with 6 antennas, and 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 the sake of simplicity and clarity in the presentation, 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[4].

A. Convergence Evaluation

In this subsection, we evaluate the convergence behavior and the computational time of the methods presented in Sections IV and V. In the figures, the legend is interpreted as follows: 'GDC' refers to the method in Section IV-A, 'GDP' to the method in Section IV-B, and 'DCP' to the method in Section V. We set the values of Q_j for the three users as $\mathbf{Q} = [3.8, 7.2, 6.4]$ power units. Software package CVX is used to solve problem 13 [15].

[4]The weights could be adjusted to assign different priorities to users [14], although this is out of the scope of this paper.

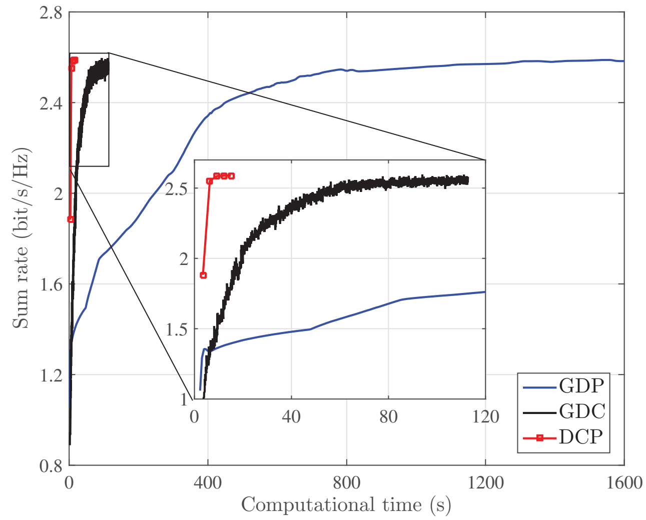

Fig. 2: Convergence of the system sum rate vs computational time of the three proposed approaches.

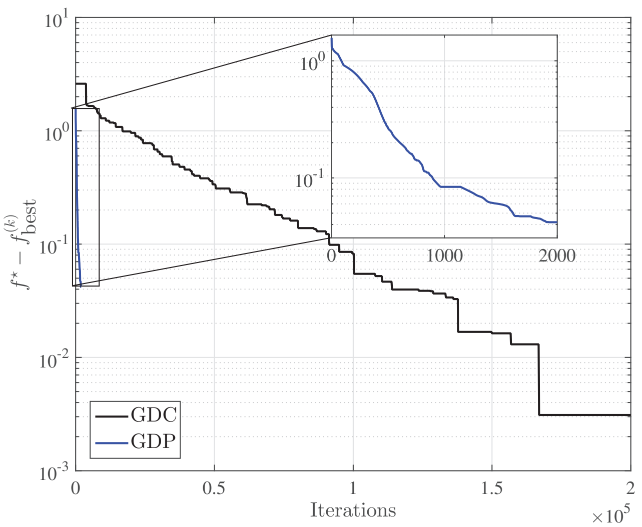

Fig. 3: Convergence of the system sum rate vs number of iterations of gradient approaches.

Fig. 2 presents the sum rate convergence as a function of computational time. The three approaches converge to the same sum rate value but requiring a different execution time. As we can see, the DC programming approach is faster than the gradient-based approaches and, in particular, much faster than the GDP approach. The GDP approach yields intermediate feasible solutions but at a expense of solving a convex problem. This involves many operations and this is reflected in the large computational time GDP requires to converge.

Fig. 3 shows the convergence behavior as a function of iterations. The DCP approach is not shown in this plot, but taking a look at Fig. 2, we see that DCP only requires 5 iterations to converge. We also see that GDP requires a lot fewer iterations than GDC but each iteration involves solving an optimization problem and, thus, the overall computational time increases a lot (see Fig. 2).

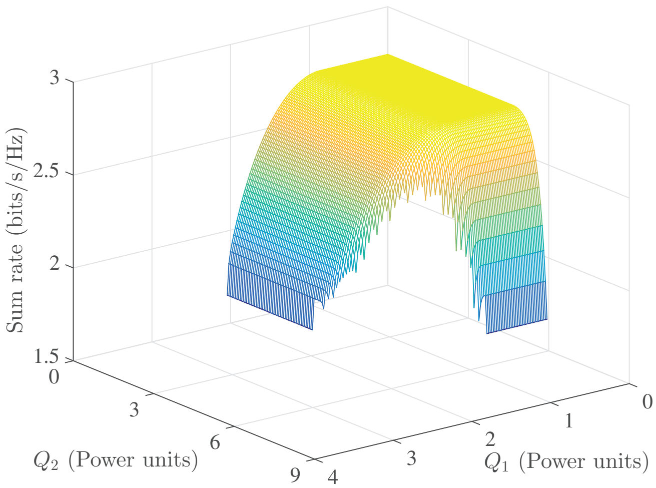

Fig. 4: Rate-Power curve for the BD method.

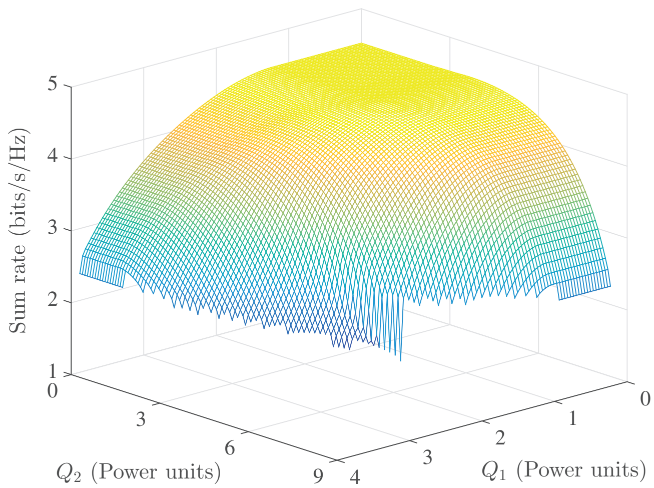

Fig. 5: Rate-Power curve for the proposed method.

B. Performance Evaluation

In this section, we evaluate the performance of the proposed approaches as compared to the classical BD strategy considered in the literature (see for example [5], [16]). In order to show how harvesting users at different distances affect the performance, we have generated channel matrices with different norms. The gradient-based approach and the DC programming approach generate the same result so both of them can be used to generate the next figures. 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. 4 and Fig. 5 show the Rate-Power region, that is, the multidimensional trade-off between the system sum rate and the power to be collected by harvesting users (see [4] for a formal definition of the Rate-Power region). As we see, the proposed approach outperforms the BD strategy in both terms, system sum rate and harvested power. The maximum system sum rate obtained with the proposed 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

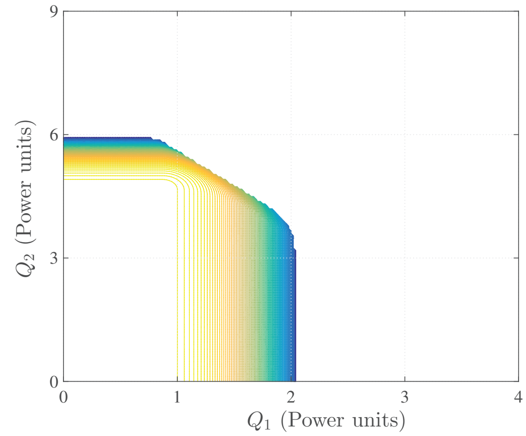

Fig. 6: Contour of Rate-Power curve for the BD method.

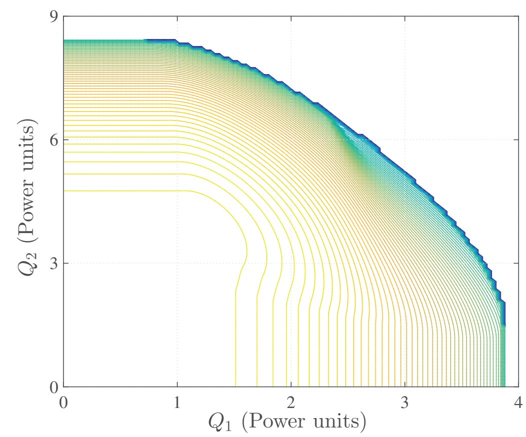

Fig. 7: Contour of Rate-Power curve for the proposed method.

generated by varying the $\{Q_j\}$ in problem (4). Note, however, that the whole Rate-Power curve need not be generated for each transmission, it is just the representation of the existing rate-power tradeoff. In order to clearly see the benefits in terms of collected power, Fig. 6 and Fig. 7 show the contour plots of the previous 3D plots. We observe that the users in the DC approach collect roughly 50% more power than the one collected by users following the BD strategy.

VII. Conclusions

In this paper, we have presented a method to solve the difficult nonconvex problem that arises in multiuser multi-stream broadcast MIMO SWIPT networks. We have formulated the general SWIPT problem as an optimization problem, in which the system weighted sum rate is optimized considering per-user harvested constraints. We have proposed two different approaches to solve the previous nonconvex problem. The first approach is based on a classical gradient-based method for constrained optimization and the second approach is based on a DC approach in which we have derived a convex approximation for the nonconvex objective. Simulation results have shown that the proposed method outperforms 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 of the DC approach required for convergence has been shown to be really low, around one to two orders of magnitude lower than the gradient-based approaches.

APPENDIX

Let us start with the gradient of the objective function. As the covariance matrices $\mathbf{S}_i$ have a particular structure (they are positive semidefinite matrices, i.e., $\mathbf{S}_i \succeq 0$, we have to follow the steps presented in [17] to obtain the desired gradient. We will first consider that matrices $\mathbf{S}_i$ are unpatterned matrices denoted by $\tilde{\mathbf{S}}_i$. Then, we will particularize the results for the specific pattern they have. Given this, differential of $f_0(\tilde{\mathbf{S}})$ (being $\tilde{\mathbf{S}} \triangleq (\tilde{\mathbf{S}}_i)_{i\in\mathcal{U}_I}$) with respect to $\tilde{\mathbf{S}}_\ell$ is

$$\begin{aligned} \mathrm{d}f_0(\tilde{\mathbf{S}}) = & \sum_{i\in\mathcal{U}_I} \omega_i \,\mathrm{Tr}\left(\left(\mathbf{H}_i \sum_{k\in\mathcal{U}_I} \tilde{\mathbf{S}}_k \mathbf{H}_i^H + \mathbf{I}\right)^{-1} \mathbf{H}_i \mathrm{d}\tilde{\mathbf{S}}_\ell \mathbf{H}_i^H\right) \\ & - \sum_{\substack{i\in\mathcal{U}_I \\ i\neq\ell}} \omega_i \,\mathrm{Tr}\left(\left(\mathbf{\Omega}_i\right)^{-1} \mathbf{H}_i \mathrm{d}\tilde{\mathbf{S}}_\ell \mathbf{H}_i^H\right), \end{aligned} \tag{14}$$

where we have used $\mathrm{d}\log\det(\mathbf{X}) = \mathrm{Tr}(\mathbf{X}^{-1}\mathrm{d}\mathbf{X})$ [17], and the gradient with respect to $\tilde{\mathbf{S}}_\ell$ and $\tilde{\mathbf{S}}_\ell^*$ are given, thus, by

$$\begin{aligned} \nabla_{\tilde{\mathbf{S}}_\ell} f_0(\tilde{\mathbf{S}}) = & \sum_{i\in\mathcal{U}_I} \omega_i \mathbf{H}_i^T \left(\mathbf{H}_i \sum_{k\in\mathcal{U}_I} \tilde{\mathbf{S}}_k \mathbf{H}_i^H + \mathbf{I}\right)^{-T} \mathbf{H}_i^* \\ & - \sum_{\substack{i\in\mathcal{U}_I \\ i\neq\ell}} \omega_i \mathbf{H}_i^T \left(\mathbf{\Omega}_i\right)^{-T} \mathbf{H}_i^*. \end{aligned} \tag{15}$$

$$\nabla_{\tilde{\mathbf{S}}_\ell^*} f_0(\tilde{\mathbf{S}}) = \mathbf{0}. \tag{16}$$

Now, for the particular case of having Hermitian matrices, the following relation holds

$$\nabla_{\mathbf{S}_\ell} f_0\left(\mathbf{S}\right) = \left[\nabla_{\tilde{\mathbf{S}}_\ell} f_0\left(\tilde{\mathbf{S}}\right) + \left(\nabla_{\tilde{\mathbf{S}}_\ell^*} f_0\left(\tilde{\mathbf{S}}\right)\right)^T\right]_{\tilde{\mathbf{S}}_\ell = \mathbf{S}_\ell}, \tag{17}$$

and, since $\mathbf{S}_\ell^* = \mathbf{S}_\ell^T$, it follows that

$$\begin{aligned} \nabla_{\mathbf{S}_\ell^*} f_0\left(\mathbf{S}\right) &= \nabla_{\mathbf{S}_\ell^T} f_0\left(\mathbf{S}\right) = \left(\nabla_{\mathbf{S}_\ell} f_0\left(\mathbf{S}\right)\right)^T && (18) \\ &= \left[\nabla_{\tilde{\mathbf{S}}_\ell^*} f_0\left(\tilde{\mathbf{S}}_i\right) + \left(\nabla_{\tilde{\mathbf{S}}_\ell} f_0\left(\tilde{\mathbf{S}}_i\right)\right)^T\right]_{\tilde{\mathbf{S}}_\ell = \mathbf{S}_\ell}. && (19) \end{aligned}$$

Finally, from (15), (16), and (19), it follows that the gradient of $f_0\left(\mathbf{S}\right)$ with respect to $\mathbf{S}_\ell^*$ is given by

$$\begin{aligned} \nabla_{\mathbf{S}_\ell^*} f_0\left(\mathbf{S}\right) = & \sum_{i\in\mathcal{U}_I} \omega_i \mathbf{H}_i^H \left(\mathbf{H}_i \sum_{k\in\mathcal{U}_I} \mathbf{S}_k \mathbf{H}_i^H + \mathbf{I}\right)^{-1} \mathbf{H}_i \\ & - \sum_{\substack{i\in\mathcal{U}_I \\ i\neq\ell}} \omega_i \mathbf{H}_i^H \left(\mathbf{\Omega}_i\right)^{-1} \mathbf{H}_i. \end{aligned} \tag{20}$$

Now we follow the same procedure for the constraints. Note that imposing that matrix $\mathbf{S}_i$ is positive semidefinite is the same as imposing that the eigenvalues of $\mathbf{S}_i$ are all non-negative. Given this, the differentials of constraints $C1, C2$, and $C3$ with respect to the unpatterned matrix $\tilde{\mathbf{S}}_\ell$ are given by

$$\mathrm{d}C1 = -\,\mathrm{Tr}\left(\mathbf{H}_j \mathrm{d}\tilde{\mathbf{S}}_\ell \mathbf{H}_j^H\right), \quad \forall j \in \mathcal{U}_E \tag{21}$$

$$\mathrm{d}C2 = \mathrm{Tr}\left(\mathrm{d}\tilde{\mathbf{S}}_\ell\right) \tag{22}$$

$$\mathrm{d}C3 = \mathrm{d}\lambda_\ell^{(k)}\left(\tilde{\mathbf{S}}_\ell\right) = -\mathbf{v}_\ell^{H(k)} \mathrm{d}\tilde{\mathbf{S}}_\ell \mathbf{v}_\ell^{(k)}, \quad \forall \ell \in \mathcal{U}_I, \forall k, \tag{23}$$

where $\lambda_\ell^{(k)}(\cdot)$ is the k-th eigenvalue of the ℓ-th covariance matrix and $\mathbf{v}_\ell^{(k)}$ is the eigenvector associated with the k-th eigenvalue. Note that we have used the identity $\mathrm{d}\lambda_i = \mathbf{v}_i^H \mathrm{d}\mathbf{X}\mathbf{v}_i$ [17]. The gradients with respect to $\tilde{\mathbf{S}}_\ell$ and $\tilde{\mathbf{S}}_\ell^*$ are:

$$\nabla_{\tilde{\mathbf{S}}_\ell} C1 = -\mathbf{H}_j^T \mathbf{H}_j^*, \quad \forall j \in \mathcal{U}_E \tag{24}$$

$$\nabla_{\tilde{\mathbf{S}}_\ell} C2 = \mathbf{I} \tag{25}$$

$$\nabla_{\tilde{\mathbf{S}}_\ell} \lambda_\ell^{(k)}\left(\mathbf{S}_\ell\right) = -\mathbf{v}_\ell^{*(k)} \mathbf{v}_\ell^{T(k)}, \quad \forall \ell \in \mathcal{U}_I, \quad \forall k \tag{26}$$

$$\nabla_{\tilde{\mathbf{S}}_\ell^*} C1 = \mathbf{0}, \quad \forall j \in \mathcal{U}_E \tag{27}$$

$$\nabla_{\tilde{\mathbf{S}}_\ell^*} C2 = \mathbf{0} \tag{28}$$

$$\nabla_{\tilde{\mathbf{S}}_\ell^*} \lambda_\ell^{(k)}\left(\mathbf{S}_\ell\right) = \mathbf{0}, \quad \forall \ell \in \mathcal{U}_I, \quad \forall k \tag{29}$$

Finally, from (25)-(29), and (19) it follows that the gradients with respect to $\mathbf{S}_\ell^*$ are given by

$$\nabla_{\mathbf{S}_\ell^*} C1 = -\mathbf{H}_j^H \mathbf{H}_j, \quad \forall j \in \mathcal{U}_E \tag{30}$$

$$\nabla_{\mathbf{S}_\ell^*} C2 = \mathbf{I} \tag{31}$$

$$\nabla_{\mathbf{S}_\ell^*} \lambda_\ell^{(k)}\left(\mathbf{S}_\ell\right) = -\mathbf{v}_\ell^{(k)} \mathbf{v}_\ell^{H(k)}, \quad \forall \ell \in \mathcal{U}_I, \quad \forall k \tag{32}$$

REFERENCES

[1] X. Lu *et al.*, "Wireless networks with RF energy harvesting: A contemporary survey," *IEEE Communications Surveys and Tutorials*, vol. 17, pp. 757–789, Secondquarter 2015.

[2] J. Paradiso and T. Starner, "Energy scavenging for mobile wireless electronics," *IEEE Computing Pervasive*, vol. 4, pp. 18–27, Jan. 2005.

[3] L. R. Varshney, "Transporting information and energy simultaneously," in *International Symposium on Information Theory*, Jul. 2008.

[4] R. Zhang and C. K. Ho, "MIMO broadcasting for simultaneous wireless information and power transfer," *IEEE Trans. on Wireless Communications*, vol. 12, pp. 1989–2001, May 2013.

[5] J. Rubio and A. Pascual-Iserte, "Simultaneous wireless information and power transfer in multiuser MIMO systems," in *IEEE Global Communications Conference (GLOBECOM)*, Dec. 2013.

[6] Q. H. Spencer *et al.*, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. on Signal Processing*, vol. 52, pp. 461–471, Feb. 2004.

[7] J. Park and B. Clerckx, "Joint wireless information and energy transfer in a K-user MIMO interference channel.," *IEEE Trans. on Wireless Comm.*, vol. 13, pp. 5781–5796, Oct. 2014.

[8] Z. Zong *et al.*, "Optimal transceiver design for SWIPT in K-user MIMO interference channels," *IEEE Trans. on Wireless Comm.*, vol. 15, pp. 430–445, Jan. 2016.

[9] J. Xu, L. Liu, and R. Zhang, "Multiuser MISO beamforming for simultaneous wireless information and power transfer," *IEEE Trans. on Signal Processing*, vol. 62, pp. 4798–4810, Sep. 2015.

[10] J. Rubio, A. Pascual-Iserte, D. P. Palomar, and A. Goldsmith, "Joint optimization of power and data transfer in multiuser MIMO systems," *Submitted to IEEE Trans. Signal Processing*. Available at https://arxiv.org/abs/1604.00434.

[11] R. Horst and N. V. Thoai, "DC programming: An overview," *Journal of Optimization Theory and Application*, vol. 103, pp. 1–43, Oct. 1999.

[12] S. Boyd and L. Vandenbergue, *Convex optimization*. Cambridge, 2004.

[13] 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 Processing*, vol. 62, pp. 641–656, Feb. 2014.

[14] M. Andrews *et al.*, "Providing quality of service over a shared wireless link," *IEEE Commununications Magazine*, vol. 39, pp. 150–154, Feb. 2001.

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

[16] R. Zhang, "Cooperative multi-cell block diagonalization with per-base-station power constraints," *IEEE Journal on Selected Areas in Comm.*, vol. 28, pp. 1435–1445, Dec. 2010.

[17] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Application in Statistics and Econometrics*. John Wiley & Sons, Inc., Essex, UK, 1988.