

# A FRAMEWORK FOR GLOBALLY OPTIMAL ENERGY-EFFICIENT RESOURCE ALLOCATION IN WIRELESS NETWORKS

Alessio Zappone\*, Emil Björnson<sup>+</sup>, Luca Sanguinetti<sup>†‡</sup>, Eduard Jorswieck\*

\*Technische Universität Dresden, Communications Laboratory, Dresden, Germany.

<sup>+</sup>Department of Electrical Engineering, ISI Linköping University, Linköping, Sweden

<sup>†</sup>Dipartimento di Ingegneria dell'Informazione, University of Pisa, Pisa, Italy

<sup>‡</sup>Large Networks and System Group (LANEAS), CentraleSupélec, Université Paris-Saclay, Gif-sur-Yvette, France

## ABSTRACT

State-of-the-art algorithms for energy-efficient power allocation in wireless networks are based on fractional programming theory, and allow to find the global maximum of the energy efficiency only in noise-limited scenarios. In interference-limited scenarios, several sub-optimal solutions have been proposed, but an efficient framework to globally maximize energy-efficient metrics is lacking. The goal of this work is to fill this gap by making use of fractional programming theory jointly with monotonic optimization. The resulting optimization framework is useful for at least two main reasons. First, it sheds light on the ultimate energy-efficiency performance of wireless networks. Second, it provides the means to benchmark the energy efficiency of state-of-the-art, but sub-optimal, solutions.

**Index Terms**— Energy Efficiency, Fractional programming, Monotonic optimization, Resource allocation, Power control.

## 1. INTRODUCTION AND RELATION TO PRIOR WORK

The percentage of the total CO<sub>2</sub>-equivalent emissions due to the information and communication technology (ICT) is estimated to be 5% [1]. While this may seem a small percentage, it will rapidly increase, due to the advent of 5G networks and the associated exponential growth of connected devices. Credited sources foresee the number of connected devices to reach 50 billions by 2020, with a data traffic increase of a factor 1000 [2]. If no countermeasures are taken, the energy demand to operate and provide such massive data rates to this massive number of devices will become unmanageable, and the resulting greenhouse gas emissions and electromagnetic pollution will exceed safety thresholds. While restricting the global ICT usage is unrealistic, a promising answer to this issue lies in optimizing the energy efficiency (EE) of ICT systems, defined as the ratio between the data rate and associated energy consumption. Given the fractional nature of the EE, the main mathematical framework adopted for EE optimization is fractional programming theory [3]. However, while fractional programming algorithms exhibit polynomial-time complexity in noise-limited systems [4], their complexity becomes prohibitive in interference-limited systems [3]. A common way to circumvent this problem is to only consider sub-optimal orthogonal or semi-orthogonal transmission schemes as well as interference cancellation techniques, to fall back into the noise-limited case. In [5, 6] multi-carrier networks are considered, and the

global energy efficiency (GEE) of the system (defined as the ratio between the sum achievable rate and the total consumed power) is optimized using orthogonal or semi-orthogonal subcarrier allocation schemes. In [7], the authors consider a multiple-antenna system and aim at maximizing the GEE when non-linear interference cancellation techniques are used. However, orthogonal interference suppression schemes inevitably result in a poor resource reutilization in multi-link networks and are thus not reasonable in large networks. Also, practical impairments (such as, for example, channel estimation errors) largely reduce the performance of interference cancellation schemes and break orthogonality. Alternative approaches employ suboptimal procedures, typically based on the use of alternating optimization techniques, as in [8], where the minimum of the individual EEs is maximized and in [9, 10] where both the maximization of GEE and of the sum of the individual EEs are considered. In [11, 12] fractional programming is used in conjunction with sequential convex optimization to develop a framework able to determine local optima of the system EE with polynomial-time complexity.

The main issue with all the cited works is that, although all proposed solutions exhibit an affordable complexity, they are not guaranteed to achieve global optimality. Indeed, a framework to obtain the global solution of EE maximization problems is currently lacking. This prevents one from gaining insight on the ultimate energy-efficient performance of wireless networks and thus from benchmarking the performance of suboptimal methods against the optimal solution. This work aims at filling these gaps by developing an optimization framework to globally maximize the EE in wireless networks. This will be done by merging fractional programming theory with monotonic optimization, which is a theory that provides algorithms to globally solve certain classes of non-convex problems [13, 14]. Monotonic optimization has been previously used to globally solve power control and scheduling problems [15], as well as to develop rate maximization schemes and beamforming techniques [16, 17]. Recent surveys of monotonic optimization applied to wireless communications are [18, 19].

## 2. SYSTEM MODEL AND PROBLEM STATEMENT

Consider a wireless network wherein  $K$  mutually interfering links are active over a communication bandwidth of  $B$  Hz. Each link includes a single-antenna transmitter node and a receiver node (possibly equipped with multiple antennas). Denoting by  $\mathbf{p}$  the  $K \times 1$  vector collecting all users' transmit powers, we assume the  $k$ -th link's signal-to-interference-plus-noise ratio (SINR)  $\gamma_k(\mathbf{p})$  takes the following general form:

$$\gamma_k(\mathbf{p}) = \frac{\alpha_k p_k}{\sigma^2 + \phi_k p_k + \sum_{i=1, i \neq k}^K p_i \beta_{i,k}} \quad (1)$$

A. Zappone and E. Jorswieck are funded by the German Research Foundation (DFG) with grant CEMRIN - ZA 747/1-3, and in the Collaborative Research Center 912 "HAEC", respectively. L. Sanguinetti is supported by the ERC Starting Grant 305123 MORE. E. Björnson is supported by the Swedish Foundation for Strategic Research (SFF).

where  $\sigma^2$  accounts for the thermal noise power at the receiver (over the bandwidth  $B$ ),  $\alpha_k$  is the  $k$ -th user's channel power gain, and  $\{\beta_{i,k}\}$  are the multi-user interference coefficients, which depend on the other users' channels as well as on global system parameters. The coefficient  $\phi_k$  models a self-interference term, proportional to the useful power, which arises in several relevant instances of communication systems [20], such as hardware-impaired networks, receivers with imperfect CSI estimation, and relay-assisted communications. By setting  $\phi_k = 0$ , the expression in (1) reduces to the canonical SINR expression in interference networks. For later convenience, observe that the  $k$ -th user's achievable rate can be expressed as  $B \log_2(1 + \gamma_k(\mathbf{p})) = q_k^+(\mathbf{p}) - q_k^-(\mathbf{p})$  with

$$q_k^+(\mathbf{p}) = B \log_2 \left( \sigma^2 + (\alpha_k + \phi_k)p_k + \sum_{i=1, i \neq k}^K p_i \beta_{i,k} \right) \quad (2)$$

$$q_k^-(\mathbf{p}) = B \log_2 \left( \sigma^2 + \phi_k p_k + \sum_{i=1, i \neq k}^K p_i \beta_{i,k} \right). \quad (3)$$

The EE (measured in bit/Joule) of link  $k$  is defined as the ratio of the achievable rate and the total consumed power (neglecting the functional dependence of  $\gamma_k$  from  $\mathbf{p}$  for notational simplicity):

$$\text{EE}_k(\mathbf{p}) = \frac{B \log_2(1 + \gamma_k)}{\mu_k p_k + \Psi_k} \quad (4)$$

wherein  $\mu_k \geq 1$  is the inverse of the power amplifier efficiency of transmitter node  $k$  and  $\Psi_k$  is the circuit power required to operate link  $k$  accounting for the dissipation in analog hardware, digital signal processing, backhaul signaling, and other overhead costs (such as cooling and power supply losses) [21]. Clearly, (4) is a link-centric (or user-centric) performance metric. A network-centric definition of EE requires to combine the individual energy efficiencies of the different links. Although different approaches have been proposed, a single definition that unarguably best represents the EE of the whole network is not available, since the different EEs are typically conflicting objectives [3, 22]. Two well-established metrics to measure the network EE are the GEE and the weighted minimum energy efficiency (WMEE), defined as [3, 4]:

$$\text{GEE}(\mathbf{p}) = \frac{\sum_{k=1}^K B \log_2(1 + \gamma_k)}{\sum_{k=1}^K \mu_k p_k + \Psi_k} \quad (5)$$

$$\text{WMEE}(\mathbf{p}) = \min_{k=1, \dots, K} w_k \frac{B \log_2(1 + \gamma_k)}{\mu_k p_k + \Psi_k}. \quad (6)$$

The GEE has the strong physical meaning of network benefit-cost ratio, in terms of global amount of reliably transmitted data and global amount of consumed energy. However, it does not depend on the individual EEs, and thus it does not allow to tune the EE of the individual links according to specific needs. Instead, the WMEE is more connected to a multi-objective optimization perspective [22], in which the individual EEs are the objectives to maximize. By varying the weights in (6) it is possible to prioritize the links that require a higher EE and to describe the complete energy-efficient Pareto boundary of the system.<sup>1</sup>

The goal of this work is to find the global solution of the following optimization problem:

$$\max_{\mathbf{p}} u(\mathbf{p}) \quad \text{s.t.} \quad \mathbf{p} \in \mathcal{P} \quad (7)$$

<sup>1</sup>For the considered system, the Pareto-boundary is the outer boundary of the  $K$ -dimensional region containing all feasible energy-efficient operating points, i.e., all feasible  $K \times 1$  vectors  $[\text{EE}_1(\mathbf{p}), \dots, \text{EE}_K(\mathbf{p})]^T$ .

with

$$\mathcal{P} = \left\{ \mathbf{p} \in \mathbb{R}_+^K; \forall k \ p_k \in [0, P_{\max,k}], B \log_2(1 + \gamma_k) \geq R_{\min,k} \right\} \quad (8)$$

wherein the objective  $u(\mathbf{p})$  can be either the GEE or the WMEE,  $P_{\max,k}$  and  $R_{\min,k}$  denote the maximum feasible transmit power and minimum acceptable rate for link  $k$ , respectively.

### 3. PROBLEM SOLUTION

The optimization problem in (7) is a fractional program. However, fractional programming algorithms exhibit polynomial-time complexity only if the numerator and denominator of the fraction are respectively concave and convex, and if the constraints are also convex [3]. Unfortunately, this requirement is not fulfilled in interference-limited networks since the presence of  $q_k^-(\mathbf{p})$  as given by (3) (which is non-zero whenever multi-user interference is present) makes the numerator non-concave. For these scenarios, fractional programs are in general NP-hard, and the conventional approach is to resort to global optimization algorithms. However, these methods operate by performing an exhaustive search of the whole feasible set, with a prohibitive computational complexity and a convergence that is only guaranteed if the functions have a limited variability (e.g. Lipschitz continuity). A powerful optimization framework to reduce such computational burden and to ensure a guaranteed convergence is monotonic optimization [13]. Roughly speaking, the basic idea is that if the objective to maximize is increasing in all optimization variables, then it is not necessary to explore the complete feasible set, but only its outer boundary. The main difficulty in applying this framework to (7) is that, unlike what happens for achievable rates, both the GEE and the WMEE are not monotonic in  $\mathbf{p}$ . As shown next, this difficulty can be overcome by an interplay of fractional programming and monotonic optimization. We start by providing some necessary preliminaries on both theories and then develop the proposed optimization framework.

#### 3.1. Fractional programming

A comprehensive overview of fractional programming applied to EE maximization in wireless networks can be found in [3].

**Definition 1 (Generalized fractional program)** Let  $\mathcal{D} \subseteq \mathbb{R}^N$  and consider the functions  $f_k : \mathcal{D} \rightarrow \mathbb{R}$  and  $g_k : \mathcal{D} \rightarrow \mathbb{R}_{++}$ , with  $k = 1, \dots, K$ . A generalized fractional program is the optimization problem defined as

$$\max_{\mathbf{x}} \min_{k=1, \dots, K} \frac{f_k(\mathbf{x})}{g_k(\mathbf{x})} \quad \text{s.t.} \quad \mathbf{x} \in \mathcal{D}. \quad (9)$$

If  $K = 1$ , then (9) reduces to the so-called single-ratio fractional program:

$$\max_{\mathbf{x}} \frac{f_1(\mathbf{x})}{g_1(\mathbf{x})} \quad \text{s.t.} \quad \mathbf{x} \in \mathcal{D}. \quad (10)$$

Since the objective function in (9) is in general not concave, standard convex optimization algorithms are not guaranteed to solve (9) and specific algorithms are required. Towards this end, we have the following key result.

**Proposition 1** [23, 24]. A vector  $\mathbf{x}^* \in \mathcal{D}$  solves (9) if and only if

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{D}} \left\{ \min_{k=1, \dots, K} [f_k(\mathbf{x}) - \lambda^* g_k(\mathbf{x})] \right\} \quad (11)$$

with  $\lambda^*$  being the unique zero of the auxiliary function  $F(\lambda)$ :

$$F(\lambda) = \max_{\mathbf{x} \in \mathcal{D}} \min_{k=1, \dots, K} \{f_k(\mathbf{x}) - \lambda g_k(\mathbf{x})\}. \quad (12)$$

This result allows solving (9) by finding the unique zero of  $F(\lambda)$ . To this end, the most widely used algorithm is the (Generalized, if  $K > 1$ ) Dinkelbach's algorithm [24, 25], reported in Algorithm 1.

---

**Algorithm 1** Generalized Dinkelbach's algorithm

---

Initialize  $\lambda_0$  with  $F(\lambda_0) \geq 0$ ,  $j = 0$ ;  
**repeat**  
  Solve the problem:  
   $\mathbf{x}_j^* = \arg \max_{\mathbf{x} \in \mathcal{D}} \{ \min_{k=1, \dots, K} [f_k(\mathbf{x}) - \lambda_j g_k(\mathbf{x})] \}$ ;  
   $\lambda_{j+1} = \min_{k=1, \dots, K} \frac{f_k(\mathbf{x}_j^*)}{g_k(\mathbf{x}_j^*)}$ ;  
   $j = j + 1$ ;  
**until**  $F(\lambda_j) > \varepsilon$

---

It can be shown that the update rule for  $\lambda$  follows Newton's method applied to the function  $F(\lambda)$ . Hence, Algorithm 1 exhibits a super-linear convergence rate, but converges to the global optimum of the corresponding instance of fractional problem only provided that (11) can be *globally* solved at each iteration. If (11) is not a convex problem, this calls for global optimization algorithms that usually require to explore the whole feasible set.

### 3.2. Monotonic optimization

Monotonic optimization provides a framework to globally solve problems that exhibit monotonicity or hidden monotonicity structures [13, 14]. Some fundamental definitions and results from monotonic optimization theory are briefly recalled henceforth.

**Definition 2 (Monotonicity in  $\mathbb{R}^N$ )** A function  $f : \mathbb{R}^N \rightarrow \mathbb{R}$  is *monotonically increasing* if  $f(\mathbf{y}) \geq f(\mathbf{x})$  when  $\mathbf{y} \succeq \mathbf{x}$ , with  $\succeq$  denoting component-wise ordering.

**Definition 3 (Hyper-rectangle in  $\mathbb{R}^N$ )** Let  $\mathbf{a}, \mathbf{b} \in \mathbb{R}^N$  with  $\mathbf{a} \preceq \mathbf{b}$ . Then, the set of all  $\mathbf{x} \in \mathbb{R}^N$  such that  $\mathbf{a} \preceq \mathbf{x} \preceq \mathbf{b}$  is a *hyper-rectangle* in  $\mathbb{R}^N$  and is denoted by  $[\mathbf{a}, \mathbf{b}]$ .

**Definition 4 (Normal and Co-normal sets)** A set  $S \subset \mathbb{R}^N$  is *normal* if  $\forall \mathbf{x} \in S$ , the hyper-rectangle  $[\mathbf{0}, \mathbf{x}]$  belongs to  $S$ . A set  $S_c \subset \mathbb{R}^N$  is *co-normal* in  $[\mathbf{0}, \mathbf{b}]$  if  $\forall \mathbf{x} \in S_c$ , then  $[\mathbf{x}, \mathbf{b}] \subset S_c$ .

A given function  $h : \mathbb{R}^N \rightarrow \mathbb{R}$  defines a normal or a co-normal set if the following results hold true:

**Proposition 2 ([13])** The set  $S = \{\mathbf{x} \in \mathbb{R}^N : h(\mathbf{x}) \leq 0\}$  is *normal* and *closed* if  $h$  is *lower semi-continuous* and *increasing*. The set  $S_c = \{\mathbf{x} \in \mathbb{R}^N : h(\mathbf{x}) \geq 0\}$  is *co-normal* and *closed* if  $h$  is *upper semi-continuous* and *increasing*.

**Definition 5 (Monotonic optimization)** A *monotonic optimization problem* in canonical form is defined as

$$\max_{\mathbf{x}} f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in S \cap S_c \quad (13)$$

wherein  $f : \mathbb{R}^N \rightarrow \mathbb{R}$  is an increasing function,  $S \subset [\mathbf{0}, \mathbf{b}]$  is a compact, normal set with nonempty interior, and  $S_c$  is a closed co-normal set in  $[\mathbf{0}, \mathbf{b}]$ .

The main result of monotonic optimization theory states that the solution to (13) lies on the upper boundary of  $S \cap S_c$  [13, Proposition 7]. Therefore, methods like the polyblock algorithm [13] and the branch-reduce-and-bound (BRB) algorithm [14] can be used to globally solve (13) by searching only on the upper boundary of the feasible set, thus drastically simplifying the problem. Nevertheless, we remark that the complexity of monotonic optimization methods

is still exponential in the problem size. However, as already observed, it is much lower than general global optimization methods which do not exploit any monotonicity structure. This makes monotonic optimization attractive for the development of a framework to benchmark any suboptimal method for solving (13).

### 3.3. GEE maximization

The GEE maximization belongs to the class of single-ratio fractional problems. So, finding its solution by the Dinkelbach's algorithm requires to solve the following auxiliary problem at iteration  $j$ :

$$\max_{\mathbf{p}} \sum_{k=1}^K B \log_2(1 + \gamma_k) - \lambda_j (\mu_k p_k + \Psi_k) \quad \text{s.t.} \quad \mathbf{p} \in \mathcal{P} \quad (14)$$

for any given positive  $\lambda_j$ , and with  $\mathcal{P}$  denoting the feasible set of Problem (7). Note that the above problem is not in convex form due to the achievable rates  $\{B \log_2(1 + \gamma_k)\}$ , which also appear in the rate constraints.

**Proposition 3** Problem (14) can be reformulated as a monotonic problem in canonical form.

**Proofsketch:** Problem (14) can be equivalently written as

$$\max_{\mathbf{p}} q^+(\mathbf{p}) - q^-(\mathbf{p}, \lambda_j) \quad \text{s.t.} \quad \mathbf{p} \in \mathcal{P} \quad (15)$$

wherein  $q^+(\mathbf{p})$  and  $q^-(\mathbf{p}, \lambda_j)$  are increasing in  $\mathbf{p}$  and given by

$$q^+(\mathbf{p}) = \sum_{k=1}^K q_k^+(\mathbf{p}), \quad q^-(\mathbf{p}, \lambda_j) = \sum_{k=1}^K q_k^-(\mathbf{p}) + \lambda_j (\mu_k p_k + \Psi_k).$$

The above reformulation is not yet a monotonic optimization problem in canonical form since the difference  $q^+(\mathbf{p}) - q^-(\mathbf{p})$  of two increasing functions is not monotonic in the sense of Definition 2. To proceed further, define  $\mathbf{p}_{\max} = [P_{\max,1}, \dots, P_{\max,K}]$  and introduce the auxiliary variable  $t = q^-(\mathbf{p}_{\max}, \lambda_j) - q^-(\mathbf{p}, \lambda_j)$ . Then, for any given  $\lambda_j$ , rewrite the auxiliary problem in (15) as

$$\max_{(t, \mathbf{p})} q^+(\mathbf{p}) + t \quad \text{s.t.} \quad (t, \mathbf{p}) \in \mathcal{P} \cap \mathcal{Q} \quad (16)$$

with

$$\mathcal{Q} = \left\{ (t, \mathbf{p}) : \begin{array}{l} 0 \leq t \leq q^-(\mathbf{p}_{\max}, \lambda_j) - q^-(\mathbf{p}, \lambda_j) \\ 0 \leq t \leq q^-(\mathbf{p}_{\max}, \lambda_j) - q^-(\mathbf{0}_K, \lambda_j) \end{array} \right\}. \quad (17)$$

As for the constraints, the minimum rate constraint functions are also not monotonic in the sense of Definition 2. However, the set of constraints  $B \log_2(1 + \gamma_k) - R_{\min,k} \geq 0$  with  $k = 1, \dots, K$ , can be equivalently rewritten as the following single constraint:

$$\min_{k=1, \dots, K} [q_k^+(\mathbf{p}) - q_k^-(\mathbf{p}) - R_{\min,k}] \geq 0 \iff \quad (18)$$

$$\min_{k=1, \dots, K} \left[ q_k^+(\mathbf{p}) - \left( \sum_{i=1}^K q_i^-(\mathbf{p}) - \sum_{i=1, i \neq k}^K q_i^-(\mathbf{p}) \right) - R_{\min,k} \right] = \quad (19)$$

$$\min_{k=1, \dots, K} \left[ \underbrace{q_k^+(\mathbf{p}) + \sum_{i=1, i \neq k}^K q_i^-(\mathbf{p})}_{\widetilde{q}^+(\mathbf{p})} - R_{\min,k} \right] - \underbrace{\sum_{i=1}^K q_i^-(\mathbf{p})}_{\widetilde{q}^-(\mathbf{p})} \geq 0 \quad (20)$$

which is the difference of the two increasing functions  $\widetilde{q}^+(\mathbf{p})$  and  $\widetilde{q}^-(\mathbf{p})$ . Similarly as above, we can thus introduce the auxiliary variable  $s$  and reformulate the problem in (16) as

$$\max_{(s, t, \mathbf{p})} q^+(\mathbf{p}) + t \quad (21)$$

$$\text{s.t.} \quad (t, \mathbf{p}) \in \mathcal{Q}, \quad 0 \leq s \leq \widetilde{q}^-(\mathbf{p}_{\max}) - \widetilde{q}^-(\mathbf{0}_K)$$

$$\widetilde{q}^-(\mathbf{p}) + s \leq \widetilde{q}^-(\mathbf{p}_{\max}), \quad \widetilde{q}^+(\mathbf{p}) + s \geq \widetilde{q}^-(\mathbf{p}_{\max})$$

Finally, it can be verified that Problem (21) fulfills Definition 5, thus being a monotonic problem in canonical form. The exact details are omitted due to space constraints but they can be found in [26].  $\square$

### 3.4. WMEE maximization

The maximization of WMEE belongs to the class of generalized fractional programs and requires to solve the following auxiliary problem at iteration  $j$ :

$$\max_{\mathbf{p}} \min_{k=1, \dots, K} q_k^+(\mathbf{p}) - q_k^-(\mathbf{p}) - \lambda_j (\mu_k p_k + \Psi_k) \quad \text{s.t. } \mathbf{p} \in \mathcal{P}. \quad (22)$$

As before, the objective function is not monotonic.

**Proposition 4** *Problem (22) can be reformulated as a monotonic problem in canonical form.*

**Proofsketch:** Letting  $\nu_k(\mathbf{p}, \lambda_j) = q_k^-(\mathbf{p}) + \lambda_j (\mu_k p_k + \Psi_k)$ , we can proceed as follows:

$$\begin{aligned} q_k^+(\mathbf{p}) - \nu_k(\mathbf{p}, \lambda_j) &= q_k^+(\mathbf{p}) - \left( \sum_{i=1}^K \nu_i(\mathbf{p}, \lambda_j) - \sum_{i=1, i \neq k}^K \nu_i(\mathbf{p}, \lambda_j) \right) \\ &= \left( q_k^+(\mathbf{p}) + \sum_{i=1, i \neq k}^K \nu_i(\mathbf{p}, \lambda_j) \right) - \sum_{i=1}^K \nu_i(\mathbf{p}, \lambda_j). \quad (23) \end{aligned}$$

By similar steps as in Section 3.3, we define  $t = \sum_{i=1}^K \nu_i(\mathbf{p}_{\max}, \lambda_j) - \sum_{i=1}^K \nu_i(\mathbf{p}, \lambda_j)$ , and reformulate (22) as

$$\max_{(t, \mathbf{p})} \min_{k=1, \dots, K} q_k^+(\mathbf{p}) + \sum_{i=1, i \neq k}^K \nu_i(\mathbf{p}, \lambda_j) + t \quad \text{s.t. } (t, \mathbf{p}) \in \mathcal{P} \cap \mathcal{Q}'$$

with

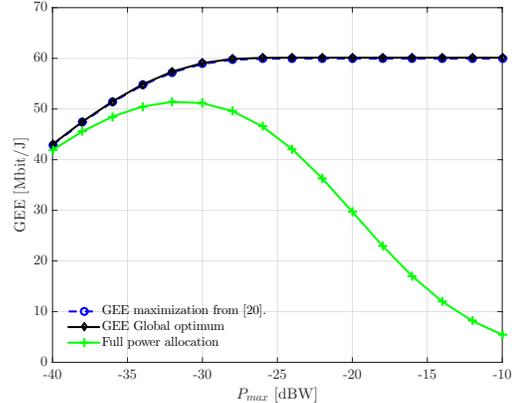
$$\mathcal{Q}' = \left\{ (t, \mathbf{p}) : \begin{array}{l} 0 \leq t \leq \sum_{i=1}^K \nu_i(\mathbf{p}_{\max}, \lambda_j) - \nu_i(\mathbf{p}, \lambda_j) \\ 0 \leq t \leq \sum_{i=1}^K \nu_i(\mathbf{p}_{\max}, \lambda_j) - \nu_i(\mathbf{0}_K, \lambda_j) \end{array} \right\} \quad (24)$$

which, reformulating the rate constraints as in Section 3.3, can be checked to fulfill Definition 5 (See [26] for more details).  $\square$

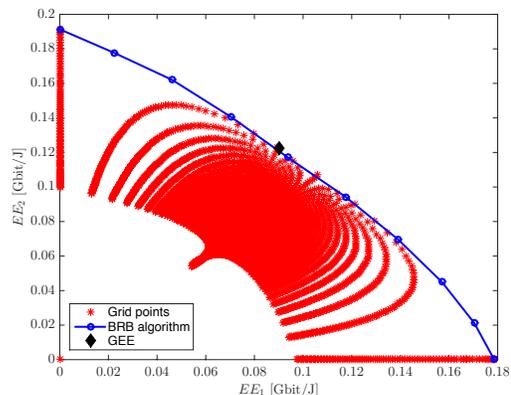
## 4. NUMERICAL RESULTS

To exemplify the benefits of the proposed optimization framework, the uplink massive MIMO scenario described in [20] is considered, with  $B = 180$  kHz. A base station equipped with 50 antennas and subject to hardware impairments serves a square area with edge of 1 km, where  $K = 5$  mobiles are randomly placed. It is shown in [20] that, assuming MRC detection, the SINR enjoyed by the generic mobile  $k$  is expressed as in (1), wherein the coefficients  $\alpha_k$ ,  $\phi_k$ , and  $\beta_{i,k}$  depend only on the propagation channels (here modeled as realizations of Rayleigh fading with path-loss model as in [27]) and on a parameter  $\epsilon \in [0, 1]$  related to the hardware quality ( $\epsilon = 0$  means ideal hardware). Here we set  $\epsilon = 10^{-2}$ . All mobiles have the same maximum feasible power  $P_{\max}$  and hardware-dissipated power  $\Psi_k = -20$  dBW. The noise power is generated as  $\sigma^2 = FB\mathcal{N}_0$ , wherein  $F = 3$  dB and  $\mathcal{N}_0 = -174$  dBm/Hz are the receiver noise figure, and noise power spectral density, respectively. No rate constraints have been enforced (i.e.  $R_{\min,k} = 0 \forall k$ ).

In Fig. 1, the GEE maximization framework is used to benchmark the (theoretically suboptimal) procedure from [20]. As a baseline scheme, we also report the GEE obtained by full power allocation (i.e.,  $p_k = P_{\max} \forall k$ ). The results are obtained by averaging over  $10^3$  independent channel scenarios. Our benchmark allows showing that the polynomial-time algorithm in [20] enjoys virtually optimal performance. Also, the achieved GEE eventually saturates,



**Fig. 1.** GEE versus  $P_{\max}$  by: (a) global optimum by monotonic optimization plus fractional programming; (b) GEE maximization algorithm proposed in [20]; (c) Full power allocation.



**Fig. 2.** Energy-efficient Pareto region for  $K = 2$ , by: (a) WMEE maximization by monotonic optimization plus fractional programming; (b) grid search over power; (c) GEE maximization point.

because for large  $P_{\max}$  only the transmit power required to achieve the global maximizer of the GEE is used. Fig. 2 plots the energy-efficient Pareto region of the system for  $K = 2$ . All other system parameters are as in Fig. 1. The Pareto region is obtained by solving the WMEE maximization problem for different choices of the weights. As a comparison, we report the energy-efficient region obtained by a grid search over the transmit powers. As seen, the proposed framework allows one to characterize the complete region, while the grid search easily fails to find all parts. Also, it is interesting to observe that the point corresponding to GEE maximization lies on the Pareto-boundary. In general, this is not guaranteed, because the GEE is not increasing in all individual EEs.

## 5. CONCLUSIONS

This work has developed an optimization framework to globally maximize the EE in wireless networks by jointly exploiting monotonic optimization theory and fractional programming theory. The framework is general enough to be applied to several instances of communication systems, such as general interference networks, massive MIMO systems, relay-assisted communications. While still exhibiting an exponential complexity, the developed framework enjoys a guaranteed convergence and a much lower complexity than standard global optimization algorithms. Also, it enables to benchmark practical algorithms, which are not theoretically guaranteed to achieve global optimality, as illustrated in the numerical results.

## 6. REFERENCES

- [1] A. Fehske, J. Malmodin, G. Biczók, and G. Fettweis, "The Global Footprint of Mobile Communications—The Ecological and Economic Perspective," *IEEE Communications Magazine, issue on Green Communications*, pp. 55–62, Aug. 2011.
- [2] Ericsson White Paper, "More than 50 billion connected devices," Ericsson, Tech. Rep. 284 23-3149 Uen, Feb. 2011.
- [3] A. Zappone and E. Jorswieck, "Energy efficiency in wireless networks via fractional programming theory," *Foundations and Trends in Communications and Information Theory*, vol. 11, no. 3-4, pp. 185–396, 2014.
- [4] C. Isheden, Z. Chong, E. Jorswieck, and G. Fettweis, "Framework for link-level energy efficiency optimization with informed transmitter," *IEEE Transactions on Wireless Communications*, vol. 11, no. 8, pp. 2946–2957, Aug. 2012.
- [5] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in multi-cell OFDMA systems with limited backhaul capacity," *IEEE Transactions on Wireless Communications*, vol. 11, no. 10, pp. 3618–3631, Oct. 2012.
- [6] Q. Xu, X. Li, H. Ji, and X. Du, "Energy-efficient resource allocation for heterogeneous services in OFDMA downlink networks: Systematic perspective," *IEEE Transactions on Vehicular Technology*, vol. 63, no. 5, pp. 2071–2082, June 2014.
- [7] J. Xu and L. Qiu, "Energy efficiency optimization for MIMO broadcast channels," *IEEE Transactions on Wireless Communications*, vol. 12, no. 2, pp. 690–701, Feb. 2013.
- [8] B. Du, C. Pan, W. Zhang, and M. Chen, "Distributed energy-efficient power optimization for CoMP systems with max-min fairness," *IEEE Communications Letters*, vol. 18, no. 6, pp. 999–1002, 2014.
- [9] S. He, Y. Huang, S. Jin, and L. Yang, "Coordinated beamforming for energy efficient transmission in multicell multiuser systems," *IEEE Transactions on Communications*, vol. 61, no. 12, pp. 4961–4971, Dec. 2013.
- [10] S. He, Y. Huang, L. Yang, and B. Ottersten, "Coordinated multicell multiuser precoding for maximizing weighted sum energy efficiency," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 741–751, Feb. 2014.
- [11] L. Venturino, A. Zappone, C. Risi, and S. Buzzi, "Energy-efficient scheduling and power allocation in downlink OFDMA networks with base station coordination," *IEEE Transactions on Wireless Communications*, vol. 14, no. 1, pp. 1–14, Jan. 2015.
- [12] A. Zappone, E. A. Jorswieck, and S. Buzzi, "Energy efficiency and interference neutralization in two-hop MIMO interference channels," *IEEE Transactions on Signal Processing*, vol. 62, no. 24, pp. 6481–6495, Dec. 2014.
- [13] H. Tuy, "Monotonic optimization," *SIAM Journal on Optimization*, vol. 11, no. 2, pp. 464–494, 2000.
- [14] H. Tuy, F. Al-Khayyal, and P. Thach, "Monotonic optimization: Branch and cut methods," in *Essays and Surveys in Global Optimization*, C. Audet, P. Hansen, and G. Savard, Eds. Springer US, 2005.
- [15] L. Qian and Y. Zhang, "S-MAPEL: Monotonic optimization for non-convex joint power control and scheduling problems," *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1708–1719, 2010.
- [16] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust monotonic optimization framework for multicell MISO systems," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2508–2523, May 2012.
- [17] W. Utschick and J. Brehmer, "Monotonic optimization framework for coordinated beamforming in multicell networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 4, pp. 1899–1909, April 2012.
- [18] E. Björnson and E. A. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Now Publishers: Foundations and Trends in Communications and Information Theory*, vol. 9, no. 2-3, pp. 113–381, Jan. 2013.
- [19] Y. J. Zhang, L. Qian, and J. Huang, "Monotonic optimization in communication and networking systems," *Now Publishers: Foundations and Trends in Networking*, vol. 7, no. 1, pp. 1–75, 2012.
- [20] A. Zappone, L. Sanguinetti, G. Bacci, E. A. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5G wireless technologies," *IEEE Transactions on Signal Processing*, vol. PP, no. 99, 2015. [Online]. Available: <http://arxiv.org/abs/1503.04609>
- [21] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3059–3075, June 2015.
- [22] E. Björnson, E. Jorswieck, M. Debbah, and B. Ottersten, "Multi-objective signal processing optimization: The way to balance conflicting metrics in 5G systems," *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 14–23, 2014.
- [23] R. Jagannathan, "On some properties of programming problems in parametric form pertaining to fractional programming," *Management Science*, vol. 12, no. 7, 1966.
- [24] W. Dinkelbach, "On nonlinear fractional programming," *Management Science*, vol. 13, no. 7, pp. 492–498, Mar. 1967.
- [25] J. P. Crouzeix and J. A. Ferland, "Algorithms for generalized fractional programming," *Mathematical Programming*, vol. 52, pp. 191–207, 1991.
- [26] A. Zappone, E. Björnson, L. Sanguinetti, and E. A. Jorswieck, "Achieving global optimality for energy efficiency maximization in wireless networks," *IEEE Transactions on Signal Processing*, 2016, under review.
- [27] G. Calcev, D. Chizhik, B. Goransson, S. Howard, H. Huang, A. Kogiantis, A. Molisch, A. Moustakas, D. Reed, and H. Xu, "A wideband spatial channel model for system-wide simulations," *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, Mar. 2007.