

MONOTONIC OPTIMIZATION FRAMEWORK FOR THE MISO IFC

Eduard A. Jorswieck

Dresden University of Technology
Communications Laboratory
01062 Dresden, Germany
eduard.jorswieck@tu-dresden.de

Erik G. Larsson

Linköping University
Dept. of Electrical Engineering (ISY)
581 83 Linköping, Sweden
erik.larsson@isy.liu.se

ABSTRACT

Resource allocation and transmit optimization for the multiple-antenna Gaussian interference channel are important but difficult problems. Recently, there has been a large interest in algorithms that find operating points which are optimal in the sum-rate, proportional-fair, or minimax sense. Finding these points entails solving a nonlinear, non-convex optimization problem. In this paper, we develop an algorithm that solves these problems exactly, to within a prescribed level of accuracy and in a finite number of steps. The main idea is to rewrite the objective functions so that methods for monotonic optimization can be used. More precisely, we write each objective function as a difference between two functions which are strictly increasing over a *normal* constraint set. The so-obtained reformulated, equivalent problem can then be solved efficiently by using so-called polyblock optimization. Numerical examples illustrate the advantages of the proposed framework compared to an exhaustive grid search.

Index Terms— Resource allocation, interference channel, non-convex optimization, outer polyblock approximation

1. INTRODUCTION

Interference channels (IFC) consist of at least two transmitters and two receivers. The first transmitter wants to transfer information to the first receiver and the second transmitter to the second receiver, respectively. This happens at the same time on the same frequency causing interference at the receivers. Information-theoretic studies of the IFC have a long history [1, 2, 3]. These references have provided various achievable rate regions, which are generally larger in the more recent papers than in the earlier ones. However, the capacity region of the general IFC remains an open problem. For certain limiting cases, for example when the interference is weak or very strong, respectively, the sum-capacity is known [4]. If the interference is weak, it can simply be treated as additional noise. For very strong interference, successive interference cancellation (SIC) can be applied at one or more of the receivers. Multiple-antenna IFCs are studied in [5]. Multiple-input multiple-output (MIMO) IFCs have also recently been studied in [6], from the perspective of spatial multiplexing gains. In [7], the rate region of the single-input single-output (SISO) IFC was characterized in terms of convexity and concavity. The MIMO IFC is also considered from a game-theoretic point of view in [8].

This work was supported in part by the Swedish Research Council (VR) and the Swedish Foundation for Strategic Research (SSF). E. Larsson is a Royal Swedish Academy of Sciences (KVA) Research Fellow supported by a grant from the Knut and Alice Wallenberg Foundation.

An explicit parameterization of the Pareto boundary for the achievable rate region of the K -user Gaussian MISO IFC, for the case when all multiuser interference is treated as additive Gaussian noise at the receivers, was derived in [9]. For the special case of two users, any point in the rate region can be achieved by choosing beamforming vectors that are linear combinations of the zero-forcing (ZF) and the maximum-ratio transmission (MRT) beamformers. Hence, all important (i.e., Pareto-efficient), operating points can be expressed by two real-valued parameters between zero and one $0 \leq \lambda = [\lambda_1, \lambda_2] \leq 1$.

In the current work, we build on the parameterization in [10] and focus on the maximum sum-rate operating point, the proportional-fair operating point and the max-min rate point. The corresponding optimization problems are non-convex problems which are difficult to solve directly. In particular, the max-min problem is non-smooth and therefore derivative-based (gradient) optimization methods cannot be applied. A suboptimal iterative algorithm based on alternating projection was proposed in [10]. In general, this algorithm converges to a local optimum. Therefore, we are interested in formulating a general non-convex optimization framework which takes as much as possible of the problem structure into account, and which is able to find the global optima of the problems.

This paper is structured as follows. First, we review the concepts of monotonic optimization and difference of monotonic functions (d.m.) maximization, and adapt these to the problem statement at hand. Next, we analyze the properties of the achievable rates as a function of λ_1 and λ_2 . The optimization problems are reformulated as difference of increasing functions programming problems, and finally, as monotonic optimization problems in a standard form. All theoretical results and the proposed algorithms are illustrated by numerical simulations. The results show the advantages of the monotonic optimization framework compared to simple exhaustive grid searches.

2. SYSTEM MODEL

In the setup that we consider, BS₁ and BS₂ have n transmit antennas each, that can be used with full phase coherency. MS₁ and MS₂, however, have a single receive antenna each. Hence our problem setup constitutes a multiple-input single-output (MISO) IFC, which is standard in the literature [5].

We assume that transmission consists of scalar coding followed by beamforming, and that all propagation channels are frequency-flat. This leads to the following basic model for the matched-filtered, symbol-sampled complex baseband data received at MS₁ and MS₂: $y_1 = \mathbf{h}_{11}^T \mathbf{w}_1 s_1 + \mathbf{h}_{21}^T \mathbf{w}_2 s_2 + e_1$, $y_2 = \mathbf{h}_{22}^T \mathbf{w}_2 s_2 + \mathbf{h}_{12}^T \mathbf{w}_1 s_1 + e_2$, where s_1 and s_2 are transmitted symbols, \mathbf{h}_{ij} is the (complex-valued) $n \times 1$ channel-vector between BS _{i} and MS _{j} , and \mathbf{w}_i is the

beamforming vector used by BS_i. The variables e_1, e_2 are noise terms which we model as i.i.d. complex Gaussian with zero mean and variance σ^2 per complex dimension. We assume that each base station can use the transmit power P , but that power cannot be traded between the base stations. Without loss of generality, we shall take $P = 1$. This gives the power constraint $\|w_i\|^2 \leq 1, i = 1, 2$. Throughout, we define the signal-to-noise ratio (SNR) as $1/\sigma^2$.

We do not consider the possibility of doing time-sharing between the systems.

3. RECENT RESULTS AND PROBLEM STATEMENT

The ZF and MRT beamformers are well known in the literature and their operational meaning in a game-theoretic framework is studied in [11]. They are given by:

$$w_1^{\text{MRT}} = \frac{h_{11}^*}{\|h_{11}\|} \quad \text{and} \quad w_2^{\text{MRT}} = \frac{h_{22}^*}{\|h_{22}\|},$$

$$w_1^{\text{ZF}} = \frac{\Pi_{h_{12}^*}^\perp h_{11}^*}{\|\Pi_{h_{12}^*}^\perp h_{11}^*\|} \quad \text{and} \quad w_2^{\text{ZF}} = \frac{\Pi_{h_{21}^*}^\perp h_{22}^*}{\|\Pi_{h_{21}^*}^\perp h_{22}^*\|}$$

for BS₁ and BS₂, respectively, where $\Pi_X^\perp = I - X(X^H X)^{-1} X^H$ denotes orthogonal projection onto the orthogonal complement of the column space of X .

The following theorem is proved in [9].

Theorem 1 Any point on the Pareto boundary of the rate region is achievable with the beamforming strategies

$$w_i(\lambda_i) = \frac{\lambda_i w_i^{\text{MRT}} + (1 - \lambda_i) w_i^{\text{ZF}}}{\|\lambda_i w_i^{\text{MRT}} + (1 - \lambda_i) w_i^{\text{ZF}}\|} \quad \text{and}$$

for some $\lambda_1, \lambda_2, 0 \leq \lambda_i \leq 1$.

The achievable rates as a function of $\lambda = [\lambda_1, \lambda_2]$ read

$$R_1(\lambda) = \log \left(1 + \frac{|w_1^T(\lambda_1) h_{11}|^2}{\sigma_n^2 + |w_2^T(\lambda_2) h_{21}|^2} \right)$$

$$R_2(\lambda) = \log \left(1 + \frac{|w_2^T(\lambda_2) h_{22}|^2}{\sigma_n^2 + |w_1^T(\lambda_1) h_{12}|^2} \right). \quad (1)$$

Based on the characterization in (1), we are interested in solving the following problems:

P1: Maximize the weighted sum-rate:

$$\max_{0 \leq \lambda \leq 1} \{\omega R_1(\lambda) + (1 - \omega) R_2(\lambda)\} \quad (2)$$

for some given $\omega, 0 \leq \omega \leq 1$.

P2: The proportional fairness problem:

$$\max_{0 \leq \lambda \leq 1} \{R_1(\lambda) \cdot R_2(\lambda)\}. \quad (3)$$

P3: The max-min problem (Egalitarian solution)

$$\max_{0 \leq \lambda \leq 1} \min\{R_1(\lambda), R_2(\lambda)\}. \quad (4)$$

All three programming problems (2), (3), and (4) are non-linear and non-convex. The iterative algorithm proposed in [10] is one possible approach to solving them, but it does not necessarily converge to the global optimum. Among algorithms that we are aware of up to this point, only an exhaustive grid search over $\lambda \in [0, 1]^2$ could

guarantee that the global optimum is found. In the following two sections, we propose a new optimization approach that finds the global solution to the problems (2), (3), and (4) to within a given accuracy and in a finite number of steps. This is our main contribution.

4. PRELIMINARIES: MONOTONIC OPTIMIZATION

Effectively the approach is to turn a non-convex but d.m. objective function (given by (2), (3) or (4)) into a strictly increasing function $\Phi(x)$. The price to pay is that we must enlarge the dimension of the problem (from 2 to 3). However, we are fortunate that the constraint set in the enlarged coefficient space is *normal* (in the sense defined in [12]). Therefore the outer polyblock approximation can be used to find the global optimum.

4.1. Increasing functions and normal sets

At first, we need the basic concepts of *increasing functions* and *normal sets*. This material is contained partly in [12]. However, we need the notion of a *strictly increasing function* and therefore we provide a complete presentation and some alternative proofs.

Definition 1 For two vectors $x', x \in \mathbb{R}^n$ we write $x' \geq x$ and say that x' dominates x if $x'_i \geq x_i$ for all $i = 1, \dots, n$. We write $x' > x$ and say that x' strictly dominates x if $x'_i > x_i$ for all $i = 1, \dots, n$.

Definition 2 A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be increasing on \mathbb{R}_+^n if $f(x) \leq f(x')$ whenever $0 \leq x \leq x'$. The function is said to be increasing in the box $[a, b]^n \subset \mathbb{R}_+^n$ if $f(x) \leq f(x')$ whenever $a \mathbf{1} \leq x \leq x' \leq b \mathbf{1}$. A function is said to be strictly increasing if for $x' \geq x \geq 0$ and $x' \neq x$ follows that $f(x') > f(x)$. (Here $\mathbf{1} = [1, \dots, 1]^T$.)

If the domain of these increasing functions is a normal set, we will later obtain a characterization of the set on which the maximum is achieved.

A set G is said to be *normal* if for all $x \in G$ all points in the box $[0, x]$ are also in G . More precisely:

Definition 3 A set $G \subset \mathbb{R}_+^n$ is called normal if for any two points $x, x' \in \mathbb{R}_+^n$ such that $x' \leq x$, if $x \in G$, then $x' \in G$, too.

For the characterization of the maximum of an increasing function over a normal set, we need the notion of *upper boundary*.

Definition 4 A point $y \in \mathbb{R}_+^n$ is called an upper boundary point of a bounded closed normal set \mathcal{D} if $y \in \mathcal{D}$ and while the set $K_y = y + \mathbb{R}_{++}^n = \{y' \in \mathbb{R}_+^n | y' > y\}$ lies outside \mathcal{D} , i.e.

$$K_y \subset \mathbb{R}_+^n \setminus \mathcal{D}.$$

The set of upper boundary points of \mathcal{D} is called the upper boundary of \mathcal{D} and it is denoted by $\partial^+ \mathcal{D}$.

In other words, a point $y \in \mathcal{D}$ is an upper boundary point of \mathcal{D} if there is no point in \mathcal{D} that strictly dominates y .

The following result shows that the maximum of a strictly increasing function over a normal set is always achieved on the upper boundary of the normal set. The statement is somewhat weaker than Proposition 7 in [12].

Proposition 1 The maximum of a strictly increasing function $f(x)$ over a normal set \mathcal{D} , if it exists, is attained on $\partial^+ \mathcal{D}$.

4.2. Monotonic optimization and polyblock approximation

The monotonic optimization problem in standard form [13] is

$$\max_{\mathbf{x}} f(\mathbf{x}) \quad \text{s.t.} \quad \mathbf{x} \in \mathcal{D} \quad (5)$$

where \mathcal{D} is a normal set. We assume that \mathcal{D} is normalized such that the smallest box containing \mathcal{D} is the unit box.

From Proposition 1 we know that the maximum of $f(\mathbf{x})$ over \mathcal{D} is attained at the upper boundary $\partial^+ \mathcal{D}$. The main idea to solve the non-convex optimization problem (5) is to approximate $\partial^+ \mathcal{D}$ by polyblocks.

Definition 5 A set $P \subset \mathbb{R}_+^n$ is called a polyblock if it is the union of a finite number of boxes.

The polyblock P is generated by a set of vertices T . The minimal set of vertices consists of only proper vertices, i.e., vertices which are not dominated by any other vertex in T . It follows that for all $z, z' \in T$ with $z \neq z'$ we have neither $z > z'$ nor $z < z'$. Another important consequence of Proposition 1 is that the maximum of an increasing function over a polyblock is achieved at a proper vertex.

The main idea of the outer polyblock algorithm is to construct a nested sequence of polyblocks $\{P_k\}$ which approximate the normal set \mathcal{D} from above, that is

$$P_1 \supset P_2 \supset \dots \supset \mathcal{D} \quad \text{such that} \\ \max\{f(\mathbf{x}) | \mathbf{x} \in P_k\} \searrow \max\{f(\mathbf{x}) | \mathbf{x} \in \mathcal{D}\}. \quad (6)$$

Define the maximizer at iteration k as

$$\tilde{\mathbf{x}}^{(k)} = \arg \max_{\mathbf{x} \in T_k} f(\mathbf{x})$$

where T_k is the minimal vertex set of P_k .

Let the set of vertices in step k be $T_k = \{\mathbf{x}_1^{(k)}, \dots, \mathbf{x}_{K^{(k)}}^{(k)}\}$. Also, let $\tilde{\mathbf{x}}^{(k)}$ denote the unique intersection point of $\partial^+ \mathcal{D}$ and $\delta \tilde{\mathbf{x}}^{(k)}$ with $\delta \in [0, 1]$. Then the set of (not necessarily minimal) vertices in step $k+1$ is constructed as follows

$$T_{k+1} = T_k \setminus \{\tilde{\mathbf{x}}^{(k)}\} \bigcup_{\nu=1}^n \{\tilde{\mathbf{x}}^{(k)} - [\tilde{x}_\nu^{(k)} - \tilde{x}_\nu^{(k)}] \mathbf{e}_\nu\} \quad (7)$$

where \mathbf{e}_n is the n th column of the identity matrix. Let P_k and P_{k+1} be the polyblocks induced by the minimal set of vertices T_k and T_{k+1} , respectively.

Proposition 2 The constructed polyblock P_k and P_{k+1} fulfill

$$\mathcal{D} \subset P_{k+1} \subset P_k \setminus \{\tilde{\mathbf{x}}^{(k)}\}. \quad (8)$$

Finally, we can remove all dominated vertices of T_{k+1} to obtain the minimal set of vertices needed for the next step $k+2$.

4.3. Outer polyblock algorithm and stopping criteria

The general outer polyblock algorithm is described in Algorithm 1. The algorithm performs two steps iteratively. First, it finds the vertex \mathbf{x} that maximizes $f(\cdot)$. Then, it subdivides the blocks in a clever way to approximate the proximity of the upper boundary point $\delta \mathbf{x} \in \partial^+ \mathcal{D}$. Next, dominated vertices are removed. The computational effort time is dominated by step that finds the intersection

Result: Solve optimization problem (5)

Input: Constraint set \mathcal{D} , accuracies ϵ and η .

```

1 initialization: Set  $T = \mathbf{1}$ ,  $k = 1$ ;
2 while  $\epsilon, \eta$ -accuracy and maximum number of steps is not
   reached do
3    $\mathbf{x}^{(k)} = \arg \max\{f(\mathbf{x}) | \mathbf{x} \in T, \mathbf{x} \geq \epsilon \mathbf{1}\}$ ;
4   if  $\mathbf{x}^{(k)} \in \mathcal{D}$  then
5      $\mathbf{x}^* = \mathbf{x}^{(k)}$  is  $\epsilon$ -optimal solution;
6   else
7     Compute the intersection point  $\mathbf{y}^{(k)}$  of  $\partial^+ \mathcal{D}$  with
        $\delta \mathbf{x}^{(k)}$  with  $0 \leq \delta \leq 1$ ;
8      $\bar{\mathbf{y}}^{(k)} = \arg \max\{f(\bar{\mathbf{y}}^{(k-1)}), f(\mathbf{y}^{(k)})\}$ ;
9     if  $f(\bar{\mathbf{y}}^{(k)}) \geq f(\mathbf{x}^{(k)}) - \eta$  then
10       $\mathbf{x}^* = \bar{\mathbf{y}}^{(k)}$  is an  $(\epsilon, \eta)$ -approximate solution of
        (5);
11    else
12      Compute  $n$  extreme points of the rectangle
         $[\mathbf{y}^{(k)}, \mathbf{x}^{(k)}]$  that are adjacent to  $\mathbf{x}^{(k)}$ :
13       $\mathbf{x}^{(k),i} = \mathbf{x}^{(k)} - (\mathbf{x}_i^{(k)} - \mathbf{y}_i^{(k)}) \mathbf{e}^i$  for  $1 \leq i \leq n$ ;
14       $Z = [T \setminus \{\mathbf{x}^{(k)}\}] \cup \{\mathbf{x}^{(k),1}, \dots, \mathbf{x}^{(k),n}\}$ ;
15       $T$  is obtained from  $Z$  after dropping all vectors
        which are dominated by others;
16    end
17  end
18   $k = k + 1$ ;
19 end
Output: Solution  $\mathbf{x}^*$  to (5)

```

Algorithm 1: General outer polyblock algorithm

point between the line to the current best vertex and the upper boundary of the constraint set. The removal of dominated vertices is efficiently implemented according to [14, Proposition 4.2]. There are three stopping criteria: when ϵ - or η -accuracy is reached, or when a maximum number of steps is exceeded.

In the implementation, we used Bolzano's bisection procedure to compute the intersection point and to determine δ in Line 7, as suggested in [13, Section 8]. Note that this problem is one-dimensional regardless of the initial problem dimension.

5. SOLUTION BY MONOTONIC OPTIMIZATION

5.1. Reformulation as d.m. problems

The next three results show that the weighted sum-rate maximization problem in (2) as well as the proportional-fair rate maximization problem in (3) and the max-min problem in (4) are d.m. programming problems.

Theorem 2 Problems P1, P2 and P3 (see Section 3) are d.m. programming problems.

Thus, the three problems of interest can be formulated as the following general d.m. problem

$$\max_{\boldsymbol{\lambda} \in [0,1]^2} \phi(\boldsymbol{\lambda}) - \psi(\boldsymbol{\lambda}) \quad (9)$$

with strictly increasing functions $\phi(\cdot)$ and $\psi(\cdot)$. Next, we substitute $\psi(\boldsymbol{\lambda}) = \psi(\mathbf{1})(1 - t)$ in (9) and obtain the equivalent programming

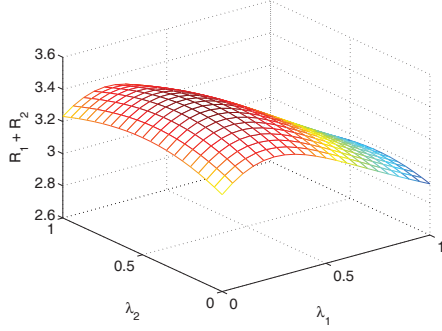


Fig. 1. Sum-rate $R_1 + R_2$ over $0 \leq \lambda \leq 1$.

problem with $\mathbf{x} = [\lambda_1, \lambda_2, t]$

$$\max_{\mathbf{x}} \underbrace{\phi(\mathbf{x}) + \psi(\mathbf{1})(x_3 - 1)}_{\Phi(\mathbf{x})} \quad \text{s.t.} \quad \mathbf{x} \in \mathcal{D} \quad (10)$$

with constraint set

$$\mathcal{D} = \{\mathbf{x} \in \mathbb{R}_+^3 : x_1 \leq 1, x_2 \leq 1, x_3 \leq 1 - \frac{\psi(x_1, x_2)}{\psi(\mathbf{1})}\}. \quad (11)$$

Note that the function $\Phi(\mathbf{x})$ is strictly increasing. The key to proceed is now:

Lemma 1 *The set \mathcal{D} defined in (11) is normal.*

Furthermore, the constraint set is compact, bounded, and connected. The programming problem in (9) corresponds exactly to the problem (5). Therefore, we can apply the outer polyblock approximation algorithm shown in Alg. 1 to solve all three problems, the weighted sum-rate maximization in (2), the proportional fair problem in (3), and the max-min problem in (4).

6. ILLUSTRATIONS

To illustrate the results, we took $n_T = 3$ and chose randomly the following channel realization:

$$\begin{aligned} \mathbf{h}_{11} &= [0.0937 + 1.1175i; 1.1264 + 0.0556i; 0.7201 + 0.4820i], \\ \mathbf{h}_{12} &= [-0.7245 + 0.3036i; -0.8728 - 0.0395i; 0.2042 + 0.2601i] \\ \mathbf{h}_{21} &= [-0.3288 - 1.4935i; 0.2623 + 0.9598i; 0.5150 + 0.7231i], \\ \mathbf{h}_{22} &= [0.7339 - 0.2231i; -0.2756 - 1.0983i; -0.9767 - 0.5006i]. \end{aligned}$$

Figure 1 shows the objective function of the problem (2) at an SNR of 0 dB. Figure 2 illustrates the upper boundary of \mathcal{D} . The function on the vertical axis $(1 - \frac{\psi(\lambda)}{\psi(\mathbf{1})})$ is non-convex, yet well approximated by the outer polyblock algorithm.

The solution found by Algorithm 1 achieves individual rates $R_1(\lambda^*) = 1.891$ and $R_2(\lambda^*) = 1.5713$ and thus a sum-rate of 3.4623. A 20×20 grid search (which corresponds to 400 function evaluations) gives the optimum as $(R_1 + R_2) = 3.4619 < (R_1(\lambda) + R_2(\lambda))$. We performed the same simulation with a 10×10 grid search and 200 polyblock iterations. The sum-rate achievable with the grid search was 3.4595 whereas the polyblock algorithm obtained a sum-rate of 3.4622. This shows the advantage of the polyblock algorithm compared to a grid search.

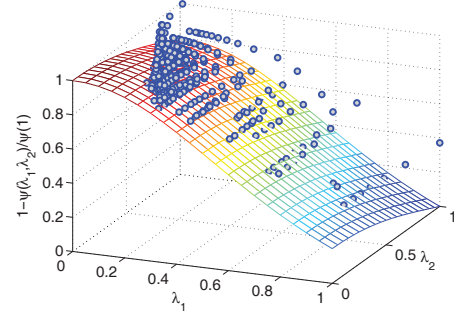


Fig. 2. Constraint set \mathcal{D} and vertices of the outer polyblock approximation after 400 iterations.

7. REFERENCES

- [1] R. Ahlswede, "The capacity region of a channel with two senders and two receivers," *Ann. Prob.*, vol. 2, pp. 805–814, 1974.
- [2] A. B. Carleial, "Interference channels," *IEEE Trans. on Inf. Theory*, vol. 24, no. 1, pp. 60–70, Jan. 1978.
- [3] T. Han and K. Kobayashi, "A new achievable rate region for the interference channel," *IEEE Trans. on Information Theory*, vol. 27, no. 1, pp. 49–60, Jan. 1981.
- [4] X. Shang, B. Chen, and M. J. Gans, "On the achievable sum rate for MIMO interference channels," *IEEE Trans. on Inf. Theory*, vol. 52, pp. 4313–4320, 2006.
- [5] S. Vishwanath and S. A. Jafar, "On the capacity of vector Gaussian interference channels," *IEEE ITW*, 2004.
- [6] S. A. Jafar and M. Fakhreddin, "Degrees of freedom for the MIMO interference channel," *IEEE Trans. on Inf. Theory*, vol. 53, pp. 2637–2642, 2007.
- [7] M. Charafeddine, A. Sezgin, and A. Paulraj, "Rate region frontiers for n-user interference channel with interference as noise," *Proc. Allerton*, 2007.
- [8] G. Scutari, D. P. Palomar, and S. Barbarossa, "Competitive design of multiuser MIMO systems based on game theory: A unified view," *IEEE Jour. on Sel. Areas in Com.*, vol. 26, pp. 1089–1103, 2008.
- [9] E. Jorswieck, E. G. Larsson and D. Danev, "Complete characterization of the Pareto boundary for the MISO interference channel," *IEEE Trans. on Sig. Proc.*, vol. 56, no. 10, Oct. 2008.
- [10] E. Jorswieck and E. G. Larsson, "The MISO interference channel from a game-theoretic perspective: A combination of selfishness and altruism achieves Pareto optimality," *Proc. of IEEE ICASSP*, 2008, invited.
- [11] E. G. Larsson and E. A. Jorswieck, "Competition versus cooperation on the MISO interference channel," *IEEE Jour. on Sel. Areas in Com.*, vol. 26, no. 7, pp. 1059–1069, Sept. 2008.
- [12] H. Tuy, "Monotonic optimization: Problems and solution approaches," *SIAM J. on Opt.*, vol. 11, no. 2, pp. 464–494, 2000.
- [13] A. Rubinov, H. Tuy, and H. Mays, "An algorithm for monotonic global optimization problems," *Optimization*, vol. 49, no. 3, pp. 205–221, 2001.
- [14] H. Tuy, F. Al-Khayyal, and P. T. Thach, *Essays and Surveys in Global Optimization*. Springer US, 2005, ch. Monotonic Optimization: Branch and Cut Methods, pp. 39–78.