# A (Very) Brief Survey on Optimization Methods for Wireless Communication Systems

A. A. P. Guimarães, I. M. Guerreiro, L. M. C. Sousa, D. C. Moreira, T. F. Maciel and C. C. Cavalcante
Wireless Telecom Research Lab (GTEL)
Universidade Federal do Ceará (UFC)
Fortaleza, Brazil
{alisson,igor,ligia,darlan,maciel,charles}@gtel.ufc.br

***Abstract*—Wireless data usage is growing now faster than ever before. In order to attend the increasing demand for wireless services and considering that frequency spectrum is a scarce and expensive resource, wireless are required to operate as efficiently as possible. In this context, the application of mathematical optimization methods in the study and design of key functionalities of wireless systems has acquired great relevance. This papers surveys some applications of optimization methods to wireless communications problems. Among them, game theory and majorization theory have got increasing attention in the last few years and are described in some more details. An application of optimization methods to solve a concrete problem in modern wireless communications, namely, the maximization of the ergodic capacity of a Coordinated Multi-Point system with statistical Channel State Information at the Transmitter is also provided.**

***Index Terms*—wireless communications, optimization methods, game theory, majorization theory.**

## I. Introduction

Wireless data usage is growing now faster than ever before. In order to attend the increasing demand for wireless services and considering that frequency spectrum is a scarce and expensive resource, modern wireless networks are required to operate as efficiently as possible.

In this context, the application of mathematical optimization methods in the study and design of key functionalities of wireless systems has acquired great relevance. The myriad of methods that find application in wireless communications is so extensive as the list of important problems that permeate this area.

The list of methods includes classic optimization methods, such as linear, convex, semi-definite, and integer optimization; multi-objective optimization tools, such as game theory; and approximative methods, such as the majorization theory and statistical approximations.

The list of problems includes power allocation, subcarrier assignment, linear and non-linear precoding, SDMA grouping, multicast beamforming, and ergodic capacity maximization.

In this work, it is not our intent to be either exhaustive or intensive in the analysis of mathematical methods applied to wireless communications, but to point out some key techniques that have found considerable application in this area.

Most applications referred in this work require a conventional background on optimization methods, covering aspects like

- optimization problem formulation: objective function(s) and problem constraints;
- problem classification: linear, quadratic, convex, concave, or semi-definite problems, among others;
- nature of optimization variables: continuous, integer, or mixed-integer together with related aspects, such as relaxations;
- multi-objective optimization;
- duality, bounds and approximations; among other aspects.

An overview of all these optimization concepts does not fit into the scope of this work. However, all these topics are widely addressed in specialized literature, e.g., in [1]–[3], among others.

This work is organized as follows. Section II provides a list of applications of optimization methods in wireless communications, specially regarding resource allocation. The applications are addressed only very briefly, more to give an rough idea of the variety of optimization problems found in wireless communications. Sections III, IV, and V address specific optimization tools which are gaining visibility in the last few years. In these sections, the topics are addressed in some more details. Section III describes some key aspects of game theory, while section IV provides the fundamentals of majorization theory. Section V illustrates an application of optimization methods to solve a concrete problem in modern wireless communications. Finally, section VI presents some conclusions.

## II. Resource Allocation in MIMO Systems

Optimization methods have found considerable application in resource allocation. A well-known example concerns the maximization of the minimum Signal to Interference-plus-Noise Ratio (SINR) among a set of co-channel links through centralized power control [4, Ch. 6], which can be reduced to an eigenvalue problem using Lagrange methods.

In more recent studies considering systems whose air interfaces base on Orthogonal Frequency Division Multiple Access (OFDMA) and Multiple Input Multiple Output (MIMO), the optimum resource allocation for a set of relevant scenarios has been determined with the help of optimization methods. For example, considering that each subcarrier $n \in \{1, 2, \dots, N\}$ of an OFDMA system is assigned to a Single-User (SU) $j \in 1, 2, \dots, J$, convex optimization can be used to show that the total system throughput is maximized by decoupling power allocation from subcarrier assignment; by assigning each subcarrier $n$ to the user $j_n^\star$ with the highest channel norm, i.e.,

$$j_n^\star = \max_j \|h_{j,n}\|_2 \quad (1)$$

with $h_{j,n}$ being the channel coefficient for user $j$ on subcarrier $n$; and by allocating power to each subcarrier using Water-Filling (WF) [5], [6]. The same result extends straightforwardly to SU MIMO (and Multiple Input Single Output (MISO) as particular case as well) systems considering linear precoding based on Singular Value Decomposition (SVD) [7], with $\|h_{j,n}\|_2$ replaced by an adequate channel norm (e.g., $\|\mathbf{H}_{j,n}\|_{\text{F}}$ or $\|\mathbf{h}_{j,n}\|_2$) and WF performed across spatial and subcarrier dimensions. Indeed, spatial precoding is fundamental for an efficient resource allocation in MIMO systems and, considering perfect Channel State Information at the Transmitter (CSIT), optimal (or near-optimal) precoding vectors can be determined with help of Lagrange and other optimization methods for several linear and nonlinear criteria, such as Matched Filter (MF), Zero-Forcing (ZF), Wiener's, and Minimum Mean Square Error (MMSE) criteria [8], [9].

Resource allocation in Multi-User (MU) MISO and MIMO scenarios also deserved considerable attention regarding optimal solutions. In [10], [11], the problem of SINR balancing with individual SINR constraints of [4] has been generalized and fast iterative algorithms for solving the problem in MU MISO scenarios have been devised by exploiting uplink-downlink duality [12]. The authors separate precoding and power optimization problems, with the former being formulated as a generalized MMSE precoding problem and solved with help of Lagrange optimization [11], [13]–[15] and the latter being formulated as a standard eigenvalue problem, as in [4], [11]. The two problems are solved alternately and the proposed algorithm is shown to converge after just a few iterations [10], [11].

While in [10], [11], as well as in other works, the set of users being multiplexed in space is predefined, the selection of these users constitutes itself an important problem, namely, the Space Division Multiple Access (SDMA) grouping problem. Since "half" users can not be selected to receive data, the SDMA grouping problem is an integer and usually non-linear, non-convex optimization problem. Several works gave optimal or suboptimal solutions to the SDMA grouping problem using different mathematical tools and heuristics, as listed in [16], [17]. For example, some authors select users according to certain optimization criteria and solve the problem using Semidefinite Programming (SDP) [18], [19]. In [20], [21], the SDMA grouping problem is formulated as an integer quadratic optimization problem, which is non-convex. By using diagonal loading, integer relaxations, and rounding, an approximation problem is formulated and solved using convex-optimization. Moreover, in [21] the subcarrier assignment to SDMA groups is formulated as an integer optimization problem, namely, a standard assignment problem, and solved using Munkres' algorithm [22].

From the above list of problems, it can be noted that different optimization methods have found many applications in wireless communications. In the sequel, some optimization tools and problems that have been acquiring increasing importance in wireless communications are described.

## III. Game Theory

### A. Introduction

Game theory is a branch of applied mathematics which provides a basis for the analysis of interactive decision-making processes [3]. It provides tools for predicting what might happen when individuals with conflicting interests interact, or more generally, for analyzing optimization problems with multiple conflicting objective functions. It also uses models to study interactions with formalized incentive structures called *games*, which are based on mathematical models of conflict and cooperation among *rational* and *intelligent* decision-makers.

An individual is said to be *rational* if each one of his decision-making behaviors is consistent with the maximization of an expected utility, and he is also said to be *intelligent* if he understands everything about the structure of the situation, including the fact that others individuals are also rational and intelligent decision-makers. In fact, these two assumptions are fundamental for the structure of the game.

Game theory can be applied in a variety of fields, including economics, international relations, evolutionary biology, political science, and military strategy.

The history of game theory originates from the works [23] by Waldegrave (1713), Cournot (1838), Darwin (1871), Edgeworth (1881), Zermelo (1913), Borel (1921), and Ville (1938). In modern times the major works by Von Neumann and Morgenstern, e.g., *Theory of Games and Economic Behavior* published in 1944, provided an axiomatic development of utility theory, which dominates the current economic thought and also introduced the formal notion of a cooperative game. Another important name is Jonh Nash, who contributed to the development of both non-cooperative and cooperative game theory [24]–[26], e.g., with the existence proof of an equilibrium in finite[1] non-cooperative games, the so-called Nash Equilibrium (NE), which is probably his most important contribution in the field.

### B. Non-cooperative Static Games Basics

In a non-cooperative game, each *player* of a set of players adjusts his strategy to optimize his own ability (utility) to compete with others. It is relatively easy to delineate the main ingredients of a conflict situation: a player has to make a decision and each possible decision leads to a different outcome or result, which are valued differently by that player. This player may not be the only one who decides about a particular outcome; a series of decisions by several individuals may be necessary. If all these players value the possible outcomes differently, the seeds for a conflict situation are there. The players involved do not always have complete control over the outcome. Indeed, some uncertainties might influence the

[1]The class of games in which the players have a *finite* number of alternatives to choose from is called finite games.

outcome in an unpredictable way so that it is (partly) based on data not yet known and not determined by the other players' decisions.

***Strategy:*** A *strategy* is a complete contingent plan, or a decision rule, that defines the *action* a player will select in every different state of the game. A simple real-world situation helps distinguishing between actions and strategies: if a player has to decide between fishing and going to work next day, then a strategy is "if the weather report predicts dry weather, then the player will go fishing, otherwise he will go to his office". Thus, what actually will be done depends on quantities not yet known and not controlled by the decision-maker, e.g., the weather condition. On the other hand, any consequence of such a strategy, after the unknown quantities are realized, is called an action. A player has a *pure-strategy* when he always picks a *single strategy* among those available in his *strategy set*. An alternative for a player is to randomize over the strategies in his strategy set, in which case the player has a *mixed-strategy*. In other words, a mixed strategy is an assignment of a probability to each pure strategy, whereas a pure strategy is selected with probability $1$ and every other strategy with probability $0$. This work will only deal with pure strategies, but further definitions can be found in [3], [27].

***Utility:*** A *utility* (payoff) function quantifies the motivations of players. A utility function for a given player assigns a (real) number for every possible outcome of the game with the property that a higher (or lower) number implies that the outcome is more (or less) preferred. Therefore, a player's strategy may be formulated as "maximizing (minimizing) his utility (cost)".

***Strategic-form Game:*** Strategic form (or normal form) is a matrix representation of a simultaneous game. For a two-player game, one player is the "row" (two-dimensional matrix) and the other is the "column'. For $Q$ players, each one is a dimension of a $Q$-dimensional matrix. Each dimension represents a strategy, and each matrix entry represents the utility to each player for every combination of strategies.

***Nash Equilibrium:*** The NE is the most common solution concept of a game. It is a joint strategy where no player can increase his utility function by unilaterally deviating [27], i.e., no player has anything to gain by changing his strategy while the other players keep their unchanged. Another NE' interpretation is that it is a mutual best response from each player to other player's strategies. It is worth mentioning that a NE is not always clearly efficient (or Pareto optimal)[2] Nevertheless, the NE remains the fundamental concept in game theory.

In [28], Nash proved that every finite strategic-form game has at least one mixed-strategy NE. As for pure-strategy NE, the uniqueness or even existence of such a NE is not guaranteed. For it, some desirable properties of the structure of a game must be established.

***Further Game Aspects:*** A game can be classified according to multiple aspects [3], [27], [29]. Some relevant types of games are described in the sequel.

[2] A Pareto-optimal solution is a joint decision of the players made in cooperation such that no other solution can improve the performance of at least one them, without degrading the performance of the other.

1) Zero-sum Games: the conflicts in a game determine a given game is classified as either *zero-sum* or *nonzero-sum*. In a zero-sum game, a gain for a player is exactly a loss for the other and the summation of the players' utility equals zero. Otherwise, the game is a nonzero-sum one.
2) Static Games: a static game is one in which all players make decisions (or select a strategy) simultaneously, without knowledge of the other players' strategies. Although decisions may be made at different points in time, the game is simultaneous because each player has no information about the decision of others.
3) Non-cooperative Games: a game is said to be non-cooperative when all players make decisions independently. Thus, while they may be able to cooperate, any cooperation must be self-enforcing. Furthermore, players can communicate with each other but cannot make a deal.

***Prisoner's Dilemma:*** A classical (and fundamental) example game is the well-known "Prisoner's Dilemma" [29], which has been popularized by the mathematician Albert W. Tucker. This example shows a hypothetical situation: two criminals, e.g., Bonnie and Clyde, are arrested for committing a crime in unison, but the police do not have enough proof to convict either. Thus, the police separate the two and offer a deal: if Bonnie testifies to convict (betrays) Clyde, she will get a sentence of 10 years if he also betrays her, or go free otherwise. However, if Bonnie does not betray (i.e., be silent) Clyde, she will get a sentence of 20 years if Clyde betrays her, or otherwise get a sentence of 2 years. The same deal is offered to Clyde. The strategic form is shown in the matrix below.

| **Bonnie** \ **Clyde** | Betray | Not Betray |
| --- | --- | --- |
| Betray | 10,10 | 0,20 |
| Not Betray | 20, 0 | 2, 2 |

Thus, each player's strategy is "Betray" or "Not Betray". This game is classified as non-cooperative and static because the players do not exchange any information with each other. Besides, they make decisions independently (i.e., separately in different rooms) and simultaneously. The only equilibrium in this example game is "Betray\Betray". However, this solution is inefficient because "Not Betray\Not Betray" provides a better output than the NE. This Pareto-optimal solution can be reached if Bonnie and Clyde cooperate.

***Application Problems:*** In modern wireless networks, signaling is normally used to obtain information, such as channel conditions, used to perform, e.g., optimal resource allocation. However, this signaling represents a considerable overhead for communications and its reduction can greatly increase spectrum utilization and the number of served users, thus improving the network performance.

One form of reducing signaling overheads is to do resource optimization using only local information. This is especially important if the system topology is distributed. In some wireless network scenarios, it is hard for an individual user to know the channel conditions of the other users. The users cannot cooperate with each other. They act selfishly to maximize their own performance in a distributed fashion.

There are some major problems in wireless communications that may fit into such model of competition for resources. Namely, we may cite power control, antenna selection, resource allocation, spectrum sharing, interference mitigation, among several other. For optimization methods in those problems we refer the reader to the following references [27], [30]–[32]. A very recent application on antenna selection is reported in [33]. A very good collection of papers dealing with game theory in wireless systems can be found on [34].

## IV. Majorization Theory

### A. Introduction

Majorization theory has been developed from the expansion of the mathematical theory of inequalities. It establishes a comparison between two vectors of $\mathbb{R}^n$ from the decreasing reordering of the their coordinates and following some restrictions. Some known results of the majorization theory include the *Lorentz curve* (1905), which determines how the distributions of income or wealth can be compared for a given population; the *principle of transfer* of Hugh Dalton (1920), which also applies to the context of income distribution; and the *Hadamard inequality* introduced by Issai Schur (1923), among other results [35].

In order to formalize these ideas, among others, the book *Inequalities* of Hardy, Littlewood and Pólya (1934) [36], was the first to unify the existing subjects about majorization theory. They have presented definitions, notations and development of results of this new mathematical formalism. After that, important results have emerged, e.g., in matrix analysis, linear algebra, optimization, and in statistical problems, which were documented in the book *Inequalities: Theory of Majorization and its Applications* by Marshall and Olkin [35], that is considered the leading reference on the subject.

The majorization theory has interesting results applied to optimization problems, such as the reformulation of non-convex problems into equivalent convexs problems [6], [37]. Consequently, majorization theory became an ally in solving problems in various subject areas, including wireless communication [37], [38]. Zhang *et al.* [39] lists some applications of the majorization theory on MIMO channels with respect to optimization problems. Another contribution concerns the study of upper and lower bounds on the ergodic channel capacity [40].

This section is dedicates to presentation and application this mathematical tool into a optimization problem studied in wireless communications. Specifically, we consider a problem involving the capacity of a MIMO channel in which the channel state information (CSI) is perfectly known at the transmitter. For this situation, we will investigate the possibility of obtaining an optimal point without the direct use of Lagrange multipliers.

### B. Majorization Theory Basics

The majorization relationship allows us to compare two vectors $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ and $\mathbf{y} = (y_1, y_2, \cdots, y_n)$ of $\mathbb{R}^n$ from its coordinates. For that purpose, we consider the vectors $[\mathbf{x}] = \big(x_{[1]},\ x_{[2]}, \cdots,\ x_{[n]}\big)$ and $[\mathbf{y}] = \big(y_{[1]},\ y_{[2]}, \cdots,\ y_{[n]}\big)$ also of $\mathbb{R}^n$ obtained by reordering the coordinates in a decreasing order from the vectors $\mathbf{x}$ and $\mathbf{y}$, respectively. Thus, $x_{[1]} \geq x_{[2]} \geq \cdots \geq x_{[n]}$ and $y_{[1]} \geq y_{[2]} \geq \cdots \geq y_{[n]}$. The vector $\mathbf{x}$ is majorized by $\mathbf{y}$, and writes $\mathbf{x} \prec \mathbf{y}$, if the following conditions are met [35]:

$$\sum_{i=1}^{k} x_{[i]} \leq \sum_{i=1}^{k} y_{[i]}, \quad 1 \leq k \leq n-1, \tag{2a}$$

$$\sum_{i=1}^{n} x_{[i]} = \sum_{i=1}^{n} y_{[i]}. \tag{2b}$$

In other words, the vector $\mathbf{y}$ majorizes the vector $\mathbf{x}$ if the coordinates of $\mathbf{y}$ are more "dispersed" or "spread out" than the coordinates of $\mathbf{x}$ [35], [37]. To the understanding of this definition, we consider the vectors $\mathbf{x} = (3;\ 3;\ 3;\ 3;\ 3)$, $\mathbf{y} = (5;\ 4;\ 3;\ 2;\ 1)$ and $\mathbf{z} = (7;\ 5;\ 2;\ 0,8;\ 0,2)$ of $\mathbb{R}^5$. Note that, we have the following majorization relationships: $\mathbf{x} \prec \mathbf{y} \prec \mathbf{z}$. Figure 1 illustrates the behavior of the coordinates of its vectors.

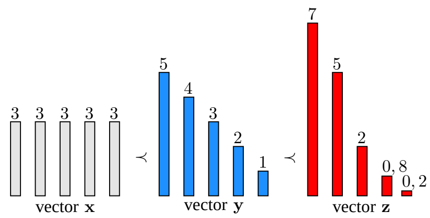



Figure 1. Geometric interpretation of the majorization relationship.

There is an extensive list of properties involving the majorization relationship, which can be found in [35], [41]. However, for this work, we highlight the following:

$$\frac{X}{n}\,\mathbf{1} \prec \mathbf{x}, \tag{3}$$

where $x_i \geq 0$, $\sum_{i=1}^{n} x_i = X$ and $\mathbf{1}$ is a vector of $\mathbb{R}^n$ whose coordinates are equal to 1, i.e., $\mathbf{1} = (1, 1, \cdots, 1)$.

In the mathematical fundamentals, a real function $f : I \subset \mathbb{R} \to \mathbb{R}$ is said to be nondecreasing in $I$ if, $\forall x_1, x_2 \in I,\ x_1 \leq x_2,\ f(x_1) \leq f(x_2)$. Similarly, Schur (1923) [35] generalized this concept of order preservation by considering the case of a real function of several variables. In this case, the domain is considered a subset of $\mathbb{R}^n$ whose elements can compared through majorization. Specifically, a real-valued function $\varphi : \mathcal{A} \subset \mathbb{R}^n \to \mathbb{R}$ is said to be *Schur-convex* on $\mathcal{A}$ if [35], [37], [38]

$$\mathbf{x} \prec \mathbf{y} \text{ on } \mathcal{A}, \text{ implies } \varphi(\mathbf{x}) \leq \varphi(\mathbf{y}). \tag{4}$$

Similarly, $\varphi(\cdot)$ is said to be *Schur-concave* on $\mathcal{A}$ if

$$\mathbf{x} \prec \mathbf{y} \text{ on } \mathcal{A}, \text{ implies } \varphi(\mathbf{y}) \leq \varphi(\mathbf{x}). \tag{5}$$

There are some criteria that allow us to specify whether a function is Schur-convex or Schur-concave, without requiring the direct application of the definition. One of them, is the following theorem [35, Proposition 3.H.2] which considers the set $\mathcal{D} = \{\mathbf{x} = (x_1, x_2, \cdots, x_n) \in \mathbb{R}^n \ ;\ x_1 \geq x_2 \geq \cdots \geq x_n\}$ whose entries are arranged in a descending sequence.

*Theorem 4.1:* Let $\varphi : \mathcal{D} \to \mathbb{R}$ be a real function defined by $\varphi(\mathbf{x}) = \sum_{i=1}^{n} g_i(x_i)$, where each $g_i : \mathbb{R} \to \mathbb{R}$ is differentiable. Then $\varphi(\cdot)$ is Schur-convex on $\mathcal{D}$ if and only if

$$g_i'(a) \geq g_{i+1}'(b) \quad \text{whenever} \quad a \geq b, \quad i = 1, 2, \cdots, n-1. \quad (6)$$

Similarly, $\varphi(\cdot)$ is Schur-concave on $\mathcal{D}$ if and only if $g_i'(a) \leq g_{i+1}'(b)$ whenever $a \geq b$.

For understanding this result, we consider the function $\varphi : \mathcal{D} \to \mathbb{R}$ defined by

$$\varphi(\mathbf{x}) = \sum_{i=1}^{n} \log_2 (1 + k\alpha_i x_i), \quad (7)$$

where, $k \geq 0$ and $\alpha_i \geq \alpha_{i+1} \geq 0$. Note that, each function $g_i(x) = \log_2 (1 + k\alpha_i x)$ is concave. In addition, $g_i'(a) \leq g_{i+1}'(b)$ whenever $a \geq b$, $i = 1, 2, \cdots, n-1$. Thus, $\varphi(\cdot)$ is Schur-concave function on $\mathcal{D}$.

Another criterion to characterize Schur-convex and Schur-concave functions is called *Schur's condition* [38, Lemma 2.5]. From this condition, we can verify also that the definitions of convexity and Schur-convexity functions are not equivalents, i.e., there is Schur-convex function that is not convex function (see [41, Example II.3.15]).

### C. Majorization Theory and Optimization

It was mentioned in the introduction of this section, the majorizaton theory is an interesting tool in solving optimization problems. The next result illustrates the applicability of this theory in a optimization problem as follow [38, Theorem 2.21]

*Theorem 4.2:* Consider the Schur-concave function $\varphi : \mathcal{D}_+ \to \mathbb{R}$ and the following optimization problem:

$$\max_{\mathbf{x} \in \mathcal{D}_+} \varphi(\mathbf{x}), \quad (8a)$$

$$\text{subject to: } \sum_{i=1}^{n} x_i = X, \quad (8b)$$

where $\mathcal{D}_+ \subset \mathcal{D}$ and $x_i \geq 0$ whenever $\mathbf{x} \in \mathcal{D}_+$. Then the global maximum is achieved by $\overline{\mathbf{x}} = \frac{X}{n}\mathbf{1}$.

*Proof:* Note that, the point $\overline{\mathbf{x}} = \frac{X}{n}\mathbf{1}$ satisfies the problem constraint. In addition, since $\varphi(\cdot)$ is a Schur-concave function and the majorization relationship in (3) is satisfied, we have $\varphi(\mathbf{x}) \leq \varphi(\overline{\mathbf{x}})$, $\forall\ \mathbf{x} \in \mathcal{D}_+$. In other words, $\overline{\mathbf{x}}$ is a optimum point. □

### D. Application Problem

We have showed in this section some results about majorization theory. In oder to apply this knowledge, we consider a single-user MIMO communication with $n_T$ transmit antennas and $n_R$ receive antennas. We assume the number of transmit antennas does not exceeds the number of receive antennas, i.e., $n_T \leq n_R$. The received signal vector $\mathbf{y} \in \mathbb{C}^{n_R \times 1}$ is given by

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \quad (9)$$

where the data vector $\mathbf{x} \in \mathbb{C}^{n_T \times 1}$ is the transmitted signal vector satisfying the total power constraint $E\left\{\|\mathbf{x}\|_2^2\right\} \leq P_T$. The channel matrix $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$ is considered deterministic and of rank $r$. The noise vector $\mathbf{n} \in \mathbb{C}^{n_R \times 1}$ is considered to be Zero-Mean Circularly Symmetric Complex Gaussian (ZMCSCG) with covariance matrix $\sigma^2\mathbf{I}_{n_R}$.

Thus, when the channel is known at the transmitter, the channel capacity is given by [7]

$$C = \sum_{i=1}^{r} \log_2 \left[1 + \frac{p_i P_T}{\sigma^2 n_T}\lambda_i\right], \quad (10)$$

where $\lambda_i$ is a eigenvalue of $\mathbf{H}\mathbf{H}^H$ with $\lambda_i \geq \lambda_{i+1}$ and $p_i$ is the power allocated in the $i$th sub-channel, which satisfies $\sum_{i=i}^{r} p_i = n_T$.

According to Paulraj *et al.* [7], the mutual information maximization problem is given by

$$\max_{\mathbf{p}} \varphi(\mathbf{p}) = \sum_{i=1}^{r} \log_2 \left[1 + \frac{p_i P_T}{\sigma^2 n_T}\lambda_i\right], \quad (11a)$$

$$\text{subject to: } \sum_{i=1}^{r} p_i = n_T. \quad (11b)$$

In addition, this is a convex optimization problem, since the objective function and constraint are concave functions. Thus, its guarantees the existence of an optimal point $\mathbf{p}^{\text{opt}} = \left(p_1^{\text{opt}}, p_2^{\text{opt}}, \cdots, p_r^{\text{opt}}\right)$ which maximizes the objective function [1]. By the Lagrange multipliers method each $p_i^{\text{opt}}$ is determined as follows [7]

$$p_i^{\text{opt}} = \left(\mu - \frac{n_T\sigma^2}{P_T\lambda_i}\right)^+, \quad i = 1, 2, \cdots, r, \quad (12)$$

where $\mu$ is the water-fill level and $(x)^+ = \max\{x, 0\}$.

If $\mathbf{p} \in \mathcal{D}_+$, then we have an alternative method which allows us to obtain an optimal point without the use of Lagrange multipliers. In fact, the function $\varphi(\cdot)$ is Schur-concave by situation presented in (7). In addition, by Theorem 4.2 the global maximum is achieved by $\mathbf{p}^{\text{opt}} = \frac{n_T}{r}\mathbf{1}$. In particular, if the channel matrix has full rank, i.e., $r = n_T$, then the optimum point is given by $\mathbf{p}^{\text{opt}} = \mathbf{1}$.

Finally, from this brief presentation of the majorization theory, we highlight the potential of the method in optimization problems. Since this mathematical tool is still incipiently studied in wireless communication problems, we visualize potential prospects of research, e.g., for broadcast and relay channels.

## V. Statistical Precoding for CoMP Systems Using Convex Optimization

Coordinated Multi-Point (CoMP) transmission/reception is a candidate technique for increasing cell-average and cell-edge throughputs in future wireless systems. In CoMP systems, several network nodes (potentially distributed over a geographic area) are linked by means of a fast backhaul to a controller unit which centrally coordinates the actions of the nodes. Joint Processing (JP) is a technique which can enhance CoMP systems' performance, mainly by employing precoding algorithms based on CSIT gathered with help of the backhaul infrastructure. In general, most precoding techniques often rely on the assumption that the transmitter knows perfectly the

MIMO channel matrix [42], [43]. However, this may not be realistic in many practical scenarios and considering partial availability of CSIT in MIMO systems becomes an important issue, which might have a significant impact on the spectral efficiency of the system.

In [44] the authors optimize the input covariance matrices in order to maximize the approximation of the ergodic capacity. Firstly, they find the ergodic capacity for a downlink MU MIMO CoMP system and evaluate a 2nd-order approximation for the ergodic capacity considering that the transmitter has access to both the mean and the covariance matrices of the channel. In the sequence, they derive the optimal input covariance using convex optimization tools.

The considered system model in [44] is the downlink of a multi-cell MU MIMO system composed by $N_b$ Base Stations (BSs) and $K$ cochannel Mobile Stations (MSs) arbitrarily distributed within the system coverage area. Each BS is equipped with $N_t$ transmit antennas and each MS with $N_r$ receive antennas. The BSs are synchronously connected to a central processing unit, thus characterizing a CoMP structure. Hereafter ($N_t$, $N_r$, $N_b$, $K$) will be used to represent the overall structure of the system. Figure 2 shows this representation for a case with $N_b = K = 3$.

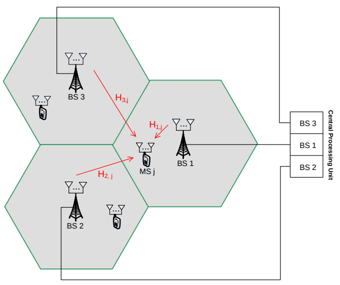



Figure 2. Multicell multiuser MIMO system model with $N_b = K = 3$.

The channel is considered frequency-flat, stationary and takes into account only Rayleigh-distributed small-scale fading, which is modeled using Jake's model [45]. The spatial channel characteristics assume Kronecker-structured covariances. $\mathbf{R}_{\mathbf{t}j}$ and $\mathbf{R}_{\mathbf{r}j}$ denote the transmit and the receive covariance matrix of MS $j$, respectively. Considering this, the channel matrix $\mathbf{H}_{\Sigma j}[n]$ from all BSs to MS $j$ at time $n$ can be modeled as

$$\mathbf{H}_{\Sigma j}[n] = \bar{\mathbf{H}}_{\Sigma j} + \mathbf{R}_{\mathbf{r}j}^{1/2}\mathbf{H}_{\text{Jakes}}[n]\mathbf{R}_{\mathbf{t}j}^{1/2}, \tag{13}$$

where $\mathbf{H}_{b,j}$ is the channel matrix from BS $b$ to MS $j$, $\mathbf{H}_{\Sigma j} = \left[\mathbf{H}_{1,j}\ \mathbf{H}_{2,j}\ \ldots\ \mathbf{H}_{N_b,j}\right]_{N_r\times N_b N_t}$ is the joint channel matrix from all BSs to MS $j$ and $\mathbf{H}_{\text{Jakes}}[n]$ is a $N_r \times N_t N_b$ small-scale fading channel matrix. Moreover, $\bar{\mathbf{H}}_{\Sigma j}$ is the mean information representing the line-of-sight component of the channel. Here, $\mathbf{R}_{\mathbf{t}j}^{1/2}$ is the principal square-root of $\mathbf{R}_{\mathbf{t}j}$, such that, $\mathbf{R}_{\mathbf{t}j} = \mathbf{R}_{\mathbf{t}j}^{1/2}\mathbf{R}_{\mathbf{t}j}^{1/2}$. Analogously, $\mathbf{R}_{\mathbf{r}j} = \mathbf{R}_{\mathbf{r}j}^{1/2}\mathbf{R}_{\mathbf{r}j}^{1/2}$.

Let $L_j$ denote the number of data streams intended for MS $j$, $j = 1, 2, \ldots, K$. For each MS, an $N_t N_b \times L_j$ precoding matrix $\mathbf{T}_j$ is designed based on the characteristics of $\mathbf{H}_{\Sigma j}$. Thus, the transmitted signal for user $j$ is $\mathbf{x}_j = \mathbf{T}_j\mathbf{s}_j$, where $\mathbf{s}_j$ is the data stream intended for MS $j$. For simplicity, streams are assumed to have i.i.d. ZMCSCG entries, i.e., Gaussian signaling is assumed.

Moreover, the transmit signal must obey a per-base power constraint given by:

$$\begin{aligned} E\left\{\sum_{k=1}^{K}\operatorname{tr}\left\{\mathbf{x}_k^{[b]}\mathbf{x}_k^{[b]H}\right\}\right\} &= \operatorname{tr}\left\{\mathbf{T}^{[b]}\mathbf{T}^{[b]H}\right\} = \\ = \operatorname{tr}\left\{\sum_{k=1}^{K}\mathbf{T}_k^{[b]}\mathbf{T}_k^{[b]H}\right\} &\leqslant P_b \quad b = 1, 2, \ldots, N_b, \end{aligned} \tag{14}$$

where $\mathbf{T}^{[b]}$ and $\mathbf{T}_k^{[b]}$ comprises the rows in $\mathbf{T}$ and $\mathbf{T}_k$ corresponding to the transmit antennas at BS $b$, respectively, and $P_b$ is the power constraint of BS $b$.

In [12], the problem of maximizing the sum rate of a MU-MIMO system is considered. Therein, uplink-downlink duality is used to transform the problem into a well-structured convex uplink MU-MIMO channel problem. In this work, the considered downlink channel has $M$ transmit antennas and each receiver has $N$ receive antennas.

### A. Ergodic Downlink Capacity Optimization of a MU MIMO CoMP System

In order to adapt the work in [12] to the scenario CoMP MU-MIMO, the authors in [44] propose some changes in the model proposed by Jindal. They use the statement shown in [12] that says that the sum capacity of the downlink MU-MIMO is equal to the sum capacity of the dual uplink MU MIMO subject to a total power constraint $P$. Therefore, they compute the sum capacity considering a global power restriction, given by the sum of the power restriction of each base, and later, to apply a power allocation matrix that satisfies the per-BS power restrictions.

Thus, the optimization problem is to maximize the sum rate capacity of the uplink MIMO CoMP system with the global power constraint, it means [44]:

$$C_{\text{CoMP}} = \max_{\{\mathbf{Q}_i\}_{i=1}^{K};\mathbf{Q}_i\geq 0;\sum_{i=1}^{K}\operatorname{tr}\{\mathbf{Q}_i\}\leq P} \log\left|\mathbf{I} + \mathbf{G}_i^H\mathbf{Q}_i\mathbf{G}_i\right|, \tag{15}$$

where $\mathbf{G}_i = \mathbf{H}_i\left(\mathbf{I} + \sum_{j\neq i}\mathbf{H}_j^H\mathbf{Q}_j\mathbf{H}_j\right)^{-1/2}$, $\sum_{b=1}^{N_b} P_b = P$, and $\mathbf{Q}_i$ is the uplink input covariance matrix per each user $i$.

Since it is not be realist in many practical scenarios to consider that the channel is perfectly known at the transmitter, in [44] it is assumed that the transmitter has access to statistical channel state information (CSI), while the receiver has access to instantaneous CSI. Thus, equation (13) is used in (15).

Using the Taylor expansion $(\log(\mathbf{I} + \mathbf{A}) = \mathbf{A} - \frac{1}{2}\mathbf{A}^2 + \frac{1}{3}\mathbf{A}^3 - \ldots)$ in (15) and considering only the two first terms, the authors in [44] obtain, after some mathematical manipulations, the 2nd-order approximation of the ergodic capacity as:

$$\begin{aligned} \bar{C}_{\text{CoMP}} = \max_{\{\mathbf{Q}_i\}_{i=1}^{K};\mathbf{Q}_i\geq 0;\sum_{i=1}^{K}\operatorname{tr}\{\mathbf{Q}_i\}\leq P} & \operatorname{tr}\{\mathbf{X}_i\mathbf{Q}_i\} - \\ & -\frac{1}{2}\left(\operatorname{tr}\left\{\mathbf{R}_{\mathbf{t}i}\mathbf{C}_i^{-1}\right\}^2\left(\operatorname{tr}\{\mathbf{R}_{\mathbf{r}i}\mathbf{Q}_i\}\right)^2 + \right. \\ & + \operatorname{tr}\{\mathbf{R}_{\mathbf{r}i}\mathbf{Q}_i\}\operatorname{Tr}\left(\bar{\mathbf{H}}_{\Sigma i}^H\mathbf{Q}_i\bar{\mathbf{H}}_{\Sigma i}\mathbf{C}_i^{-1}\mathbf{R}_{\mathbf{t}i}\mathbf{C}_i^{-1}\right) + \\ & \left.\operatorname{tr}\{\mathbf{X}_i\mathbf{Q}_i\}^2\right), \end{aligned} \tag{16}$$

being

$$\mathbf{X}_i = \left(\bar{\mathbf{H}}_{\Sigma i}\mathbf{C}_i^{-1}\bar{\mathbf{H}}_{\Sigma i}^H + \mathrm{tr}\left\{\mathbf{R}_{\mathbf{t}i}\mathbf{C}_i^{-1}\right\}\mathbf{R}_{\mathbf{r}i}\right). \tag{17}$$

Thus, the optimization problem which must be solved is:

$$\max_{\{\mathbf{Q}_i\}_{i=1}^K;\mathbf{Q}_i\geq 0;\sum_{i=1}^K \mathrm{tr}\{\mathbf{Q}_i\}\leq P} F(\mathbf{Q}_i), \tag{18}$$

where

$$\begin{aligned} F(\mathbf{Q}_i) = & \frac{1}{\delta+1}\mathrm{tr}\left\{\mathbf{X}_i\mathbf{Q}_i\right\} - \frac{1}{2(\delta+1)^2}\Big(\mathrm{tr}\left\{\mathbf{R}_{\mathbf{t}i}\mathbf{C}_i^{-1}\right\}^2 \\ & \left(\mathrm{tr}\left\{\mathbf{R}_{\mathbf{r}i}\mathbf{Q}_i\right\}\right)^2 + \mathrm{tr}\left\{\mathbf{R}_{\mathbf{r}i}\mathbf{Q}_i\right\} \\ & \mathrm{Tr}\left(\bar{\mathbf{H}}_{\Sigma i}^H\mathbf{Q}_i\bar{\mathbf{H}}_{\Sigma i}\mathbf{C}_i^{-1}\mathbf{R}_{\mathbf{t}i}\mathbf{C}_i^{-1}\right) + \mathrm{tr}\left\{\mathbf{X}_i\mathbf{Q}_i\right\}^2\Big), \end{aligned} \tag{19}$$

$\mathbf{X}_i$ is defined in (17), and $\delta$ is a parameter which guarantee the convergence and must be appropriately chosen for the SNR of interest.

Using the Karush-Kuhn-Tucker (KKT) conditions to attain both the primal and dual optimal solutions [1], the Lagrangian of (19) can be written as [44]:

$$\begin{aligned} \mathbf{L}(\mathbf{Q}_i,\mathbf{Z}_i,\nu) = & -F(\mathbf{Q}_i) - \mathrm{tr}\left\{\mathbf{Z}_i\mathbf{Q}_i\right\} + \nu(\mathrm{tr}\left\{\mathbf{Q}_i - P\right\}) \\ = & \frac{1}{2(\delta+1)^2}\Big(\mathrm{tr}\left\{\mathbf{R}_{\mathbf{t}i}\mathbf{C}_i^{-1}\right\}^2\left(\mathrm{tr}\left\{\mathbf{R}_{\mathbf{r}i}\mathbf{Q}_i\right\}\right)^2 + \\ & + \mathrm{tr}\left\{\mathbf{R}_{\mathbf{r}i}\mathbf{Q}_i\right\}\mathrm{tr}\left\{\bar{\mathbf{H}}_{\Sigma i}^H\mathbf{Q}_i\bar{\mathbf{H}}_{\Sigma i}\mathbf{C}_i^{-1}\mathbf{R}_{\mathbf{t}i}\mathbf{C}_i^{-1}\right\} + \\ & + \mathrm{tr}\left\{\mathbf{X}_i\mathbf{Q}_i\right\}^2\Big) - \frac{1}{\delta+1}\mathrm{tr}\left\{\mathbf{X}_i\mathbf{Q}_i\right\} - \\ & - \mathrm{tr}\left\{\mathbf{Z}_i\mathbf{Q}_i\right\} + \nu(\mathrm{tr}\left\{\mathbf{Q}_i - P\right\}) \end{aligned} \tag{20}$$

where $\mathbf{Z}_i$ and $\nu$ are dual variables.

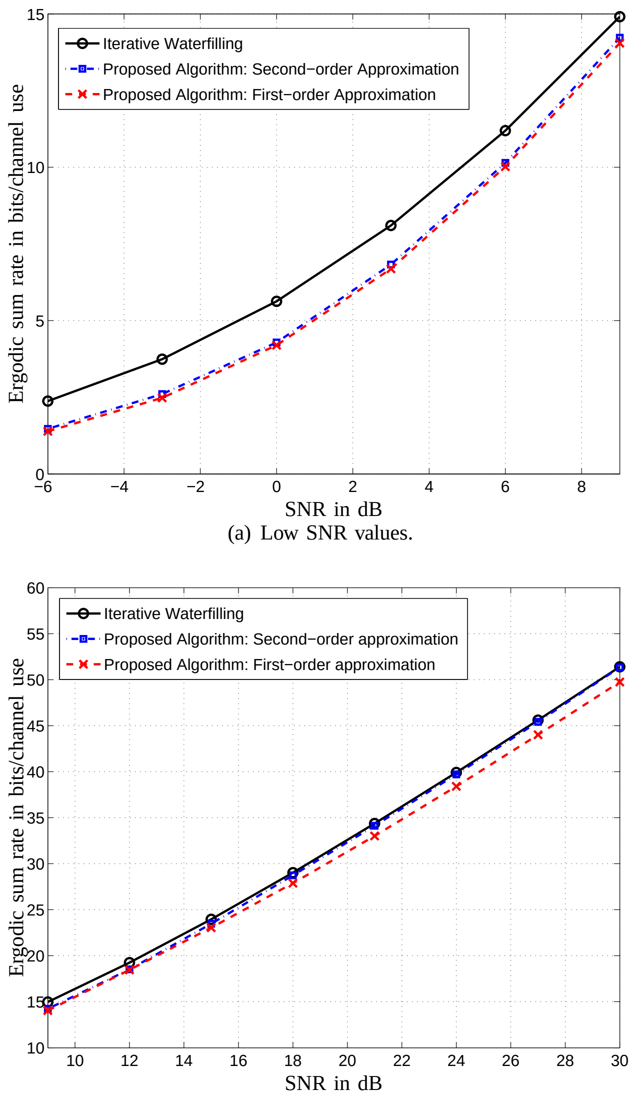
If $\mathbf{X}_i$ is invertible, the matrix $\mathbf{Q}_i$ which maximizes $F(\mathbf{Q}_i)$ is given by [44]:

$$\begin{aligned} \mathbf{Q}_i = & (\delta+1)^2\mathbf{X}_i^{-1/2}\Big(\tilde{\mathbf{Z}}_i + \frac{1}{\delta+1}\mathbf{I} + \nu\mathbf{X}^{-1} - \\ & - \Theta_1\tilde{\mathbf{R}}_{\mathbf{r}i} - \Theta_2\tilde{\mathbf{S}}_i\Big)\mathbf{X}_i^{-1/2}, \end{aligned} \tag{21}$$

where

$$\begin{aligned} \tilde{\mathbf{Z}}_i = & \ \mathbf{X}_i^{1/2}\mathbf{Z}_i\mathbf{X}_i^{1/2}, \quad \tilde{\mathbf{R}}_{\mathbf{r}i} = \mathbf{X}_i^{1/2}\mathbf{R}_{\mathbf{r}i}\mathbf{X}_i^{1/2}, \\ \tilde{\mathbf{S}}_i = & \ \mathbf{X}_i^{1/2}\bar{\mathbf{H}}_{\Sigma i}\mathbf{C}_i^{-1}\mathbf{R}_{\mathbf{t}i}\mathbf{C}_i^{-1}\bar{\mathbf{H}}_{\Sigma i}^H\mathbf{X}_i^{1/2}, \\ \Theta_1 = & \ (\delta+1)^2\left(\mathrm{tr}\left\{\mathbf{R}_{\mathbf{t}i}\mathbf{C}_i^{-1}\right\}^2\mathrm{tr}\left\{\mathbf{R}_{\mathbf{r}i}\mathbf{Q}_i\right\} + \right. \\ & \left. + \mathrm{tr}\left\{\bar{\mathbf{H}}_{\Sigma i}\mathbf{C}_i^{-1}\mathbf{R}_{\mathbf{t}i}\mathbf{C}_i^{-1}\bar{\mathbf{H}}_{\Sigma i}^H\right\}\right), \\ \Theta_2 = & \ (\delta+1)^2\,\mathrm{tr}\left\{\mathbf{R}_{\mathbf{r}i}\mathbf{Q}_i\right\}. \end{aligned}$$

In [44], the author compare the average sum rate obtained with the proposed algorithm and the upper bound on the ergodic sum rate, which corresponds to the perfect CSIT and performs an iterative water-filling [12] when the number of transmit and receive antennas ($N_t$ and $N_r$) is equal to 2. Such result can be seen in Figure 3. Moreover, They consider the case where the ergodic sum rate is approximated by only the first term of the Taylor expansion. We can notice that in low SNRs, the proposed algorithms in [44] have similar performance and the difference gap between them and the upper bound is high. When the SNR increases, the performance of the proposed algorithms improves and the $2^{\text{nd}}$-order algorithm performs close to the upper bound. In such simulations, they assume the CoMP-cell scenario consisting of 3 coordinated cells with BSs placed in the center of each cell and one user is placed randomly in each cell. These users are considered as the ones selected by a scheduling algorithm to transmit.




(a) Low SNR values.

(b) High SNR values.

Figure 3. Comparison of the ergodic sum rate obtained with the proposed algorithms and with the upper bound (iterative water-filling) in a scenario (2,2,3,3).

Finally, we can conclude that the use of convex optimization tools in the design of input covariance matrix obtains good simulation results. Thus, convex optimization tools have key functionalities in solving problems of the wireless systems.

## VI. Conclusions

In this paper, we provided a very brief survey on optimization methods applied to solve relevant problems in wireless communications. As exposed, there is a large number of optimization methods, ranging from classic mathematical to more distributed multi-objective optimization. A considerable (but not complete at all) list of references in the areas have been provided. Moreover, a concrete application of optimization to solve a wireless optimization problem has also been described.

## REFERENCES


[1] S. Boyd and L. Vandenberghe, *Convex optimization*, 1st ed. Cambridge Univ. Press, 2004.

[2] G. Nemhauser and L. Wosley, *Integer and combinatorial optimization*. Wiley & Sons, 1999.

[3] T. Basar and G. J. Olsder, *Dynamic noncooperative game theory*, 2nd ed. Society for Industrial and Applied Mathematics (SIAM), 1998.

[4] J. Zander, S.-L. Kim, M. Almgren, and O. Queseth, *Radio Resource Management for Wireless Networks*, 1st ed., ser. Mobile Communications. London, England: Artech House, 2001.

[5] H. Liu and G. Li, *OFDM-based broadband wireless networks - design and optimization*, 1st ed. Wiley & Sons, 2005.

[6] D. P. Palomar, M. A. Lagunas, and J. M. Cioffi, "Optimum linear joint transmit-receive processing for MIMO channels with QoS constraints," *IEEE Trans. Signal Process.*, vol. 52, no. 5, pp. 1179–1197, may 2004.

[7] A. Paulraj and R. Nabar and D. Gore, *Introduction to Space-Time Wireless Communications*. Cambridge University Press, 2003.

[8] M. Joham, "Optimization of linear and nonlinear transmit signal processing," Ph.D. dissertation, Munich University of Technology, Munich, Germany, Jun. 2004.

[9] M. Joham, W. Utschick, and J. A. Nossek, "Linear transmit processing in MIMO communications systems," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.

[10] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18–28, Jan. 2004.

[11] ——, *QoS-based resource allocation and transceiver optimization*, 1st ed. now publishers, Jun. 2006.

[12] N. Jindal, W. Rhee, S. Vishwanath, S. Jafar, and A. Goldsmith, "Sum Power Iterative Water-filling for Multi-antenna Gaussian Broadcast Channels," *IEEE Trans. on Information Theory*, vol. 51, no. 4, pp. 1570–1580, April 2005.

[13] P. Zetterberg, "A comparison of two systems for down link communication with antennas at the base," in *Proc. of the IEEE Wireless Commun. System Symposium*, Nov. 1995, pp. 15–20.

[14] ——, "A comparison of two systems for downlink communication with base station antenna arrays," *IEEE Trans. Veh. Technol.*, vol. 48, no. 5, pp. 1356–1370, Sept. 1999.

[15] A. B. Gershman, *Smart antennas: state of the art*. Hindawi Publish. Coorp., 2005, ch. Uplink robust adaptive beamforming, pp. 173–188.

[16] T. F. Maciel, "Suboptimal resource allocation for multi-user MIMO-OFDMA systems," Ph.D. dissertation, Technische Universität Darmstadt, Darmstadt, Germany, Sept. 2008.

[17] T. F. Maciel and A. Klein, "On the performance, complexity, and fairness of suboptimal resource allocation for multiuser MIMO-OFDMA systems," *IEEE Trans. Veh. Technol.*, vol. 59, no. 1, pp. 406–419, Jan. 2010.

[18] N. D. Sidiropoulos and Z.-Q. Luo, "A semidefinite relaxation approach to MIMO detection for high-order QAM constellations," *IEEE Signal Process. Lett.*, vol. 13, no. 9, pp. 525–528, Sept. 2006.

[19] E. Matskani, N. D. Sidiropoulos, Z.-Q. Luo, and L. Tassiulas, "Joint multiuser downlink beamforming and admission control: a semidefinite relaxation approach," in *Proc. of the IEEE Internat. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, vol. 3, no. 585-588, Apr. 2007.

[20] T. F. Maciel and A. Klein, "A convex quadratic SDMA grouping algorithm based on spatial correlation," in *Proc. of the IEEE Internat. Conf. on Commun. (ICC)*, Jun. 2007.

[21] ——, "Resource allocation strategies for SDMA/OFDMA systems," *Lecture Notes in Electrical Engineering (Advances in mobile and wireless communications - views of the 16th IST Mobile and Wireless Communications Summit)*, vol. 16, pp. 115–132, May 2008.

[22] F. Bourgeois and J.-C. Lassalle, "An extension of the Munkres algorithm for the assignment problem to rectangular matrices," *Communications of the ACM*, vol. 14, no. 12, pp. 802–804, Dec. 1971.

[23] Walker, P., "An outline of the history of game theory," University of Canterbury, Department of Economics, Working Papers 9504. [Online]. Available: http://ideas.repec.org/p/wuk/cantec/9504.html

[24] Nash, John F., "Equilibrium points in N-person games," in *Proceedings of the National Academy of Sciences of the United States of America*, 1950.

[25] ——, "The bargaining problem," *Econometrica*, vol. 18, pp. 155–162, 1950. [Online]. Available: http://jmvidal.cse.sc.edu/library/nash50a.pdf

[26] ——, "Two-person cooperative games," *Econometrica*, vol. 21, no. 1, pp. 128–140, 1953. [Online]. Available: http://dx.doi.org/10.2307/1906951

[27] MacKenzie, A. B. and DaSilva, L. A., *Game Theory for Wireless Engineers*. USA: Morgan & Claypool, 2006.

[28] Nash, John F., "Non-cooperative games," *The Annals of Mathematics*, vol. 54, no. 2, pp. 286–295, September 1951. [Online]. Available: http://dx.doi.org/10.2307/1969529

[29] Han, Zhu and Liu, K. J. Ray, *Resource Allocation for Wireless Networks: Basics, Techniques, and Applications*. New York, NY, USA: Cambridge University Press, 2008.

[30] Scutari, G., Palomar, D. P., and Barbarossa, S., "Optimal Linear Precoding Strategies for Wideband Noncooperative Systems Based on Game Theory - Part I: Nash Equilibria," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1230–1249, March 2008.

[31] ——, "Optimal linear precoding strategies for wideband non-cooperative systems based on game theory - part II: Algorithms," *IEEE Transactions on Signal Processing*, vol. 56, no. 3, pp. 1250–1267, March 2008.

[32] ——, "Competitive design of multiuser MIMO systems based on game theory: A unified view," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1089–1103, September 2008.

[33] I. M. Guerreiro and C. C. Cavalcante, "A Distributed Approach for Antenna Subset Selection in MIMO Systems," in *The Seventh International Symposium on Wireless Communication Systems*, York, UK, September 2009.

[34] E. Jorswieck, E. Larsson, M. Luise, and H. V. Poor, Eds., *Game Theory in Signal Processing and Communications*, vol. 26, no. 5. IEEE Signal Processing Magazine, September 2009.

[35] A. W. Marshall and I. Olkin, *Theory of Majorization and Its Applications*. Academic Press, 1979.

[36] G. H. Hardy, L. E. Littlewood, and G. Pólya, *Inequalities*. Cambridge University Press, 1934.

[37] D. P. Palomar and Y. Jiang, "MIMO Transceiver Design via Majorization Theory," *Found. Trends Commun. Inf. Theory*, vol. 3, no. 4-5, pp. 331–551, 2006.

[38] E. Jorswieck and H. Boche, "Majorization and Matrix-Monotone Functions in Wireless Communications," *Found. Trends Commun. Inf. Theory*, vol. 3, no. 6, pp. 553–701, 2007.

[39] X. Zhang, D. P. Palomar, and B. Ottersten, "Statistically Robust Design of Linear MIMO Transceivers," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3678–3689, aug 2008.

[40] C. Zhong, K.-K. Wong, and S. Jin, "Capacity Bounds for MIMO Nakagami-$m$ Fading Channels," *IEEE Trans. Signal Process.*, vol. 57, no. 9, pp. 3613–3623, sep 2009.

[41] R. Bhatia, *Matrix Analisys*. Springer, 1997.

[42] G. Dietl and G. Bauch, "Linear Precoding in the Downlink of Limited Feedback Multiuser MIMO," *IEEE GLOBECOM'07*, pp. 4359–4364, Nov. 2007.

[43] V. Stankovic and M. Haardt, "Generalized design of multi-user MIMO precoding matrices," *IEEE Trans. on Wireless Comm.*, vol. 7, no. 3, pp. 953–961, Mar. 2008.

[44] L. Sousa, A. Guimaraes, T. Maciel, and C. Cavalcante, "Ergodic Sum-Rate Maximization with Statistical CSIT in a Cooperative MIMO System," Eurasip Journal on Advances in Signal Processing-Special Issue on Cooperative MIMO Multicell Networks (submitted to).

[45] W. Jakes, *Microwave Mobile Communications*, 2nd ed. Piscataway, NJ: Wiley-IEEE Press, May 1994.