Optimal Transmit Beamforming Using Convex Optimization

Mats Bengtsson, Björn Ottersten

October 1999

IR–S3–SB–9955

Submitted to IEEE Transactions on Communications

ROYAL INSTITUTE OF TECHNOLOGY
Department of Signals, Sensors & Systems
Signal Processing
SE-100 44 STOCKHOLM

KUNGL TEKNISKA HÖGSKOLAN
Institutionen för Signaler, Sensorer & System
Signalbehandling
100 44 STOCKHOLM

Optimal Transmit Beamforming Using Convex Optimization

Mats Bengtsson, Björn Ottersten
Department of Signals, Sensors, and Systems
Royal Institute of Technology
S-100 44 Stockholm, Sweden
Email: matsb@s3.kth.se
Phone: +46 8 790 84 63
Fax: +46 8 790 72 60

Abstract— **When using antenna arrays at the base stations of a cellular system, one critical aspect is the transmission strategy. An optimal choice of beamformers, including power control, for simultaneous transmission to several co-channel users must be solved jointly for all users and base stations in an area. We formulate an optimal transmit strategy and show how the optimal solution can be calculated efficiently using interior point methods for semidefinite optimization, even though the original problem is non-linear and non-convex. The algorithm minimizes the total transmitted power under certain constraints to guarantee a specific quality of service. The method provides large flexibility in the choice of constraints and can be extended to be robust against channel uncertainties.**

Keywords— **Array signal processing, Space division multiplexing, Land mobile radio cellular systems, Power control, Optimization methods**

I. Introduction

The use of antenna arrays brings new possibilities in the design of mobile communications systems. It is well known that the system capacity is more limited in the downlink than in the uplink [27], [9]. Yet, the literature on beamforming for transmission is relatively small, compared to the well investigated topic of beamforming for a receiving antenna array. Some results on downlink beamforming can be found in [10], [26], [27], [13], [19], [14].

The simplest transmit strategy is to use standard beamforming, i.e., to point the main lobe of the antenna array in the direction of the specific receiver [15]. Knowing the channel to nearby co-channel users, it is possible to actively suppress the signal to the interfered users. In [18], Rashid-Farrokhi et al. formulated the beamforming design as a constrained optimization problem and presented an algorithm that finds a feasible solution. Later, the same authors [19] and Visotsky and Madhow [25] have independently shown that a minor modification of the original algorithm will find the global optimum with geometric convergence.

In this paper, we consider the same optimization problem, namely to minimize the total transmitted power while maintaining a certain quality of service for all users, but present an alternative solution using convex optimization. This gives several advantages. First of all, the optimal solution can be efficiently calculated using standard algorithms for semidefinite optimization [24], [21], which also quickly detect infeasible situations. Secondly, with this technique, it is very easy to introduce modifications, for example adding extra constraints on the dynamic range or adding increased robustness against channel estimation errors. With the original problem formulation, our method gives a solution in the form of a standard fixed beamformer. When modifications are introduced this will not always be true, but in these cases the optimal solution can be implemented using a time-varying beamformer – a space time code [22]. Note however that the problem formulations we consider here will not necessarily give any coding gain.

Convex optimization, in particular semidefinite programming, is one of the major breakthroughs in the area of mathematical optimization during the last decade. Semidefinite programming extends traditional linear programming by also including constraints that matrices formed by linear combinations of the problem variables are positive semidefinite. Several interior point methods, originally introduced for linear programming, have been extended to semidefinite problems and can in fact find the optimum with the same polynomial worst-case complexity as for linear problems [17]. Semidefinite programming has successfully been applied in several application areas, including combinatorial optimization [2] and control theory [24]. To our knowledge, convex optimization for beamformer design has previously only been used with traditional filter design-type constraints on pass bands and side-lobe levels [11], but not within the framework of statistically optimal beamforming.

When the antenna array is used as a receiver, i.e. in uplink mode, the instantaneous channel can be estimated directly from the received data, whereas in the downlink, the transmitting beamformer must be based on information collected in the uplink. Several schemes have been proposed for the transformation from uplink to downlink. In a Time Division Duplex (TDD) system with sufficiently short duplex distance and limited mobility, the downlink channel is virtually identical to the uplink channel, whereas in a Frequency Division Duplex (FDD) system, the channel fades independently at the two duplex frequencies. However, a statistical model of the downlink channel can be obtained from the collected uplink data using a physical model [28], [4] or model-free techniques [10], [3]. Throughout this paper, we assume that a stochastic characterization of the downlink transmission channel is known at the base station for all co-channel mobiles within its range (possible located in neighboring cells).

We study the joint design of the beamformers for all co-channel users within a large region which could include several cells. The optimal solution must be solved jointly since every transmitted signal will affect all receivers. In a practical system, a decentralized suboptimal solution may be preferable to reduce the overhead of collecting all channel measurements. The optimal solution does still provide a valuable benchmark for evaluation of other algorithms and for use in system simulations and capacity studies.

The paper is organized as follows. The system model and the basic assumptions are presented in Section II. In Section III, we first state the problem and show how it can be solved using a semidefinite relaxation. Then we show how to incorporate additional constraints and give a short overview of some suboptimal solutions. Section IV gives some remarks on the problem of combined uplink beamforming and power control and finally we present a few numerical examples in Section V. The appendix gives the details of some results used in Section III.

II. System Model

We consider a system where a number of co-channel mobile users are served by one or more base stations and each base station is equipped with an antenna array. We will design the beamformers assuming a stationary scenario where the fast (Rayleigh) fading is described by its second order properties. We also assume narrow-band signals without any time dispersion, i.e., the channel fading is frequency flat. The model can easily be extended to frequency selective channels, taking both co-channel interference and inter-symbol interference into account, see [19].

In the baseband, the signal received by the ith mobile, $r_i(t)$, $i = 1, \ldots, d$, is given by

$$r_i(t) = \sum_{k=1}^{K} \mathbf{v}_{i,k}^* \mathbf{x}_k(t) + n_i(t) \ , \tag{1}$$

where $(\cdot)^*$ denotes Hermitian vector transpose. Here, $\mathbf{x}_k(t)$ is the complex valued $m \times 1$ vector of the baseband signals transmitted at the m antenna elements of base station k, $n_i(t)$ is zero mean white complex noise with variance σ_i^2. The channel from base station k to mobile i is given by the random complex valued vector $\mathbf{v}_{i,k}$ with correlation matrix

$$\mathbf{R}_{i,k} = \mathrm{E}[\mathbf{v}_{i,k}\mathbf{v}_{i,k}^*] \ . \tag{2}$$

In the special case of line of sight transmission or in a system capable of tracking the fast fading of the downlink channel, $\mathbf{R}_{i,k} = \mathbf{h}_{i,k}\mathbf{h}_{i,k}^*$, where $\mathbf{h}_{i,k}$ is a deterministic array response vector, but in general, $\mathbf{R}_{i,k}$ can have any rank because of specular or diffuse multi-path propagation.

Each mobile is allocated to one base station and $\kappa(i)$ is used to denote the base station allocated for mobile i. Likewise, $\mathcal{I}(k) = \{i|\kappa(i) = k\}$ denotes the indices of the set of mobiles allocated to base k. We will use the shorthand notation $\mathbf{R}_i$ for $\mathbf{R}_{i,\kappa(i)}$.

The signal transmitted at base k is given by

$$\mathbf{x}_k(t) = \sum_{i \in \mathcal{I}(k)} \mathbf{w}_i s_i(t) \ , \tag{3}$$

where $s_i(t)$ is the scalar data sequence intended for user i and $\mathbf{w}_i$ is the beamforming weight vector for transmission from base $\kappa(i)$ to mobile i. For simplicity we assume that all $s_i(t)$ are uncorrelated and have the same power $\mathrm{E}[|s_i(t)|^2] = 1$.

III. Algorithms

We consider the design of the beamformers $\mathbf{w}_i$ given estimates of all $\mathbf{R}_{i,k}$ and σ_i^2. The goal is to minimize to total transmitted power for all base stations

$$P = \sum_{i=1}^{d} \|\mathbf{w}_i\|^2 \tag{4}$$

while maintaining an acceptable quality of service for all users. We will first consider an SINR threshold and then show how additional constraints can be used to gain increased robustness against estimation errors in the channel covariance matrices or to handle a limited dynamic range.

A. *Optimal Beamforming*

We consider the same problem formulation as in [19], [25], namely to minimize the total transmitted power under the constraint that the received SINR at each mobile is above a certain threshold, $\mathrm{SINR}_i \geq \gamma_i$. This gives the following optimization problem

$$\begin{aligned} &\min \sum_{i=1}^{d} \|\mathbf{w}_i\|^2 \\ &\text{s.t. } \frac{\mathbf{w}_i^* \mathbf{R}_i \mathbf{w}_i}{\sum_{n \neq i} \mathbf{w}_n^* \mathbf{R}_{i,\kappa(n)} \mathbf{w}_n + \sigma_i^2} \geq \gamma_i, \ i = 1, \ldots, d \ , \end{aligned} \tag{5}$$

or equivalently,

$$\begin{aligned} &\min \sum_{i=1}^{d} \mathbf{w}_i^* \mathbf{w}_i \\ &\text{s.t. } \mathbf{w}_i^* \mathbf{R}_i \mathbf{w}_i - \gamma_i \sum_{n \neq i} \mathbf{w}_n^* \mathbf{R}_{i,\kappa(n)} \mathbf{w}_n \geq \gamma_i \sigma_i^2, \ i = 1, \ldots, d \ . \end{aligned} \tag{6}$$

It is easy to show that all constraints must be active at the optimum [25], thus the inequality in (6) can be replaced by an equality.

We first consider the special case when all the channel matrices $\mathbf{R}_i$ to the desired mobiles have rank one. This is the case specifically treated in [19], [25] even though their algorithm can easily be extended to our more general data model. Then, $\mathbf{w}_i^* \mathbf{R}_i \mathbf{w}_i = |\mathbf{w}_i^* \mathbf{h}_i|^2$ and since the phase of the beamformer was left unspecified in (5), we can without loss of generality add the constraints $\mathbf{w}_i^*\mathbf{h}_i \geq 0$ which gives the following problem.

$$\begin{aligned} &\min \sum_{i=1}^{d} \mathbf{w}_i^* \mathbf{w}_i \\ &\text{s.t. } (\mathbf{w}_i^*\mathbf{h}_i)^2 \geq \gamma_i \left(\sum_{n \neq i} \mathbf{w}_n^* \mathbf{R}_{i,\kappa(n)} \mathbf{w}_n + \sigma_i^2 \right), \\ &\quad \mathbf{w}_i^*\mathbf{h}_i \geq 0,\ i = 1, \ldots, d\ . \end{aligned} \tag{7}$$

However, these constraints are just an affine transformation of the convex second-order cone $\{\mathbf{x}, y|\ \|\mathbf{x}\|^2 \leq y^2, y \geq 0\}$, which shows that this is a convex problem which can be efficiently solved using standard methods for convex optimization [12].

In the general case, the original constraint set, (6) is not convex, but we will show that the problem still can be efficiently solved using convex optimization.

To this end, introduce the matrices $\mathbf{W}_i = \mathbf{w}_i\mathbf{w}_i^*$ and use the rule $\mathbf{w}^*\mathbf{R}\mathbf{w} = \mathrm{Tr}[\mathbf{R}\mathbf{w}\mathbf{w}^*]$ to rewrite the problem into

$$\begin{aligned} &\min \sum_{i=1}^{d} \mathrm{Tr}[\mathbf{W}_i] \\ &\text{s.t. } \mathrm{Tr}[\mathbf{R}_i\mathbf{W}_i] - \gamma_i \sum_{n \neq i} \mathrm{Tr}[\mathbf{R}_{i,\kappa(n)}\mathbf{W}_n] = \gamma_i\sigma_i^2, \\ &\quad \mathbf{W}_i = \mathbf{W}_i^*, \\ &\quad \mathbf{W}_i \succeq 0,\ i = 1, \ldots, d\ . \end{aligned} \tag{8}$$

Here, the notation $\mathbf{W} \succeq 0$ means that $\mathbf{W}$ is positive semidefinite. Note that with the additional constraints $\mathrm{rank}[\mathbf{W}_i] = 1$, (8) is equivalent to (6), thus if the optimal solution of (8) has $\mathrm{rank}[\mathbf{W}_i] = 1$ for all i, then it is also an optimal solution of (5). Relaxing the rank of $\mathbf{W}_i$ gives a semidefinite optimization problem with an optimal cost that always is a lower bound for the original problem. This technique is called a Lagrangian relaxation [2], [24] since it is the Lagrangian dual of the dual of the original problem. For this specific problem, however, we can show a much stronger result, namely.

Theorem III.1: If (8) has a feasible solution, then it has at least one optimal solution where $\mathrm{rank}[\mathbf{W}_i] = 1$, for all $i = 1, \ldots, d$.

Proof: See the appendix. □

Thus, the optimal solution of (5) can be actually be calculated by solving the relaxation (8).

If the relaxation does not have a unique minimum point, there is no guarantee that all optimal solutions $\mathbf{W}_i$ have rank one. One counterexample is given by $\sigma_i^2 = 1$, $\gamma_i = 1/2$, $\kappa(i) = 1$ and

$$\mathbf{R}_{11} = \begin{bmatrix} 1&0&0 \\ 0&1&0 \\ 0&0&1 \end{bmatrix} \quad \mathbf{R}_{21} = \begin{bmatrix} 1&0&0 \\ 0&1&0 \\ 0&0&0 \end{bmatrix} \quad \mathbf{R}_{31} = \begin{bmatrix} 1&0&0 \\ 0&0&0 \\ 0&0&1 \end{bmatrix},$$

which has several minima with the same cost, including

$$\mathbf{W}_1 = \begin{bmatrix} 0&0&0 \\ 0&\frac{4}{3}&0 \\ 0&0&0 \end{bmatrix} \quad \mathbf{W}_2 = \begin{bmatrix} 0&0&0 \\ 0&\frac{7}{6}&0 \\ 0&0&0 \end{bmatrix} \quad \mathbf{W}_3 = \begin{bmatrix} 0&0&0 \\ 0&0&0 \\ 0&0&\frac{1}{2} \end{bmatrix}$$

but also

$$\mathbf{W}_1 = \begin{bmatrix} 0&0&0 \\ 0&\frac{2}{3}&0 \\ 0&0&\frac{2}{3} \end{bmatrix} \quad \mathbf{W}_2 = \begin{bmatrix} 0&0&0 \\ 0&\frac{5}{6}&0 \\ 0&0&0 \end{bmatrix} \quad \mathbf{W}_3 = \begin{bmatrix} 0&0&0 \\ 0&0&0 \\ 0&0&\frac{5}{6} \end{bmatrix}.$$

Thus, there is no guarantee that an algorithm for solving semidefinite problems will give the desired rank one solution. However, in practice these degenerate cases almost never occur and if the algorithm gives a high rank solution, a small perturbation can be added to one of the $\mathbf{R}_i$ matrices which will change the problem into having only a rank one solution. An alternative approach to construct a rank one optimal solution from any given optimal solution, is given in the appendix.

To conclude, the convex relaxation does always provide an optimal solution to the original problem (5).

B. Additional Constraints

With this relaxation technique, it is easy to change the constraints or add more constraints. In [5] we considered an alternative formulation of the quality of service with separate constraints on the signal to noise and signal to interference ratios, respectively. With the constraints $\mathrm{SNR}_i \geq \mu_i$ and $\mathrm{SIR}_i \geq \gamma_i$, the resulting semidefinite relaxation is given by

$$\begin{aligned} &\min \sum_{i=1}^{d} \mathrm{Tr}[\mathbf{W}_i] \\ &\text{s.t. } \mathrm{Tr}[\mathbf{R}_i\mathbf{W}_i] \geq \mu_i\sigma_i^2 \\ &\quad \mathrm{Tr}[\mathbf{R}_i\mathbf{W}_i] \geq \gamma_i \sum_{n \neq i} \mathrm{Tr}[\mathbf{R}_{i,\kappa(n)}\mathbf{W}_n] \\ &\quad \mathbf{W}_i = \mathbf{W}_i^*, \\ &\quad \mathbf{W}_i \succeq 0,\ i = 1, \ldots, d\ . \end{aligned} \tag{9}$$

We could also add constraints on the dynamic range of the power amplifier at each antenna element. The total transmitted power of element l of array k is

$$\mathrm{E}[|[\mathbf{x}_k]_l|^2] = \sum_{i \in \mathcal{I}(k)} |[\mathbf{w}_i]_l|^2 = \sum_{i \in \mathcal{I}(k)} [\mathbf{W}_i]_{ll}$$

Thus upper and lower bounds μ_{kl}^U and μ_{kl}^L, respectively, on the average power of an antenna element can be formulated by constraints of the form

$$\mu_{kl}^L \leq \sum_{i \in \mathcal{I}(k)} [\mathbf{W}_i]_{ll} \leq \mu_{kl}^U\ .$$

Similarly, we could set limits on the relative dynamic range of a single element in comparison to the total power for the whole array by replacing μ_{kl} with $\mu_{kl} \sum_{i \in \mathcal{I}(k)} \mathrm{Tr}[\mathbf{W}_i]$.

Since all these constraints are linear in the elements of $\mathbf{W}_i$, the resulting problem will still be semidefinite. However, we can not in general expect the optimal solutions to have rank one. For system evaluations and simulations, the optimum provides a lower bound for the problem, but a high rank solution can actually be implemented also in a practical system. If we allow for time varying beamformers

$\mathbf{w}_i(t)$, $\mathbf{W}_i$ can be interpreted as the correlation matrix of the beamformer, $\mathbf{W}_i = \mathrm{E}[\mathbf{w}_i \mathbf{w}_i^*]$. Thus, one possible implementation is to use a random sequence of vectors with covariance $\mathbf{W}_i$, as a time varying beamformer. Compare to [16] where a similar interpretation of $\mathbf{W}_i$ is used to evaluate the usefulness of space-time coding for different scenarios.

C. *Increased Robustness*

A common problem in connection with optimal uplink beamforming is signal cancellation caused by estimation errors in the channel covariances [7]. We could expect similar problems in the downlink, so it is interesting to introduce a design strategy that is robust against small errors in $\mathbf{R}_{ik}$. Assume that the true channel has a covariance matrix $\hat{\mathbf{R}}_{ik} + \Delta_{ik}$, where $\hat{\mathbf{R}}_{ik}$ is the available estimate. The goal is to find $\mathbf{w}_i$ such that the SINR constraints in (5) hold for all choices of Δ_{ik} with $\|\Delta_{ik}\| \le \epsilon_{ik}$ for some matrix norm $\|\cdot\|$. Since

$$\max_{\|\Delta\| \le \epsilon} |\mathrm{Tr}[\mathbf{X}\Delta]| = \epsilon \|\mathbf{X}\|_* ,$$

where $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$. The resulting problem can be formulated as

$$\begin{aligned}
&\min \sum_{i=1}^{d} \mathrm{Tr}[\mathbf{W}_i] \\
&\text{s.t. } \mathrm{Tr}[\hat{\mathbf{R}}_i \mathbf{W}_i] - \gamma_i \sum_{n \ne i} \mathrm{Tr}[\hat{\mathbf{R}}_{i,\kappa(n)} \mathbf{W}_n] \\
&\qquad \ge \gamma_i \sigma_i^2 + \epsilon_i \|\mathbf{W}_i\|_* + \gamma_i \sum_{n \ne i} \epsilon_{i,\kappa(n)} \|\mathbf{W}_n\|_*, \\
&\quad \mathbf{W}_i = \mathbf{W}_i^*, \\
&\quad \mathbf{W}_i \succeq 0,\ i = 1, \dots, d .
\end{aligned} \tag{10}$$

If we use the spectral norm of Δ_{ik}, i.e. put a limit of the maximum eigenvalue $\lambda_{\max}(\Delta_{ik}) \le \epsilon_{ik}$, then the dual norm is the absolute sum of eigenvalues $\|\mathbf{W}\|_* = \sum |\lambda_k(\mathbf{W})|$ and since all $\mathbf{W}_i$ are positive semidefinite, the optimal robust design is solved by the following semidefinite problem

$$\begin{aligned}
&\min \sum_{i=1}^{d} \mathrm{Tr}[\mathbf{W}_i] \\
&\text{s.t. } \mathrm{Tr}[\hat{\mathbf{R}}_i \mathbf{W}_i] - \gamma_i \sum_{n \ne i} \mathrm{Tr}[\hat{\mathbf{R}}_{i,\kappa(n)} \mathbf{W}_n] \\
&\qquad \ge \gamma_i \sigma_i^2 + \epsilon_i \,\mathrm{Tr}[\mathbf{W}_i] + \gamma_i \sum_{n \ne i} \epsilon_{i,\kappa(n)} \,\mathrm{Tr}[\mathbf{W}_n], \\
&\quad \mathbf{W}_i = \mathbf{W}_i^*, \\
&\quad \mathbf{W}_i \succeq 0,\ i = 1, \dots, d .
\end{aligned} \tag{11}$$

D. *Suboptimal Solutions*

Even if an efficient algorithm is available for the joint design of all beamformers, a decentralized algorithm using only local information could be preferable in a running network to decrease the overhead communication between different base stations.

A heuristic approach is to determine each beamformer separately, keeping the received SNR at the mobile of interest above the threshold μ_i and the total transmitted power to the interfered users below the threshold ξ_i,

$$\mathbf{w}_i^* \Big(\sum_{n \ne i} \mathbf{R}_{n,\kappa(i)} \Big) \mathbf{w}_i \triangleq \mathbf{w}_i^* \mathbf{Q}_i \mathbf{w}_i \le \xi_i .$$

Minimizing the transmitted power gives the following optimization problem,

$$\begin{aligned}
\mathbf{w}_i = \arg\min P_i &= \|\mathbf{w}_i\|^2 \\
\text{s.t. } \mathbf{w}_i^* \mathbf{R}_i \mathbf{w}_i &\ge \mu_i \sigma_i^2 \\
\mathbf{w}_i^* \mathbf{Q}_i \mathbf{w}_i &\le \xi_i .
\end{aligned} \tag{12}$$

The same relaxation as in the previous sections gives the semidefinite optimization problem,

$$\begin{aligned}
\mathbf{W}_i = \arg\min &\,\mathrm{Tr}[\mathbf{W}_i] \\
\text{s.t. } &\mathrm{Tr}[\mathbf{W}_i \mathbf{R}_i] \ge \mu_i \sigma_i^2 \\
&\mathrm{Tr}[\mathbf{W}_i \mathbf{Q}_i] \le \xi_i \\
&\mathbf{W}_i^* = \mathbf{W}_i \\
&\mathbf{W}_i \succeq 0 .
\end{aligned}$$

Also for this problem it is possible to show that there always exists an optimal solution $\mathbf{W}_i$ of rank one, so also (12) can be solved using semidefinite programming.

A related strategy, which has been suggested in several references including [10], [28], [7], is to use

$$\mathbf{w}_i = \arg\max \frac{\mathbf{w}_i^* \mathbf{R}_i \mathbf{w}_i}{\mathbf{w}_i^* (\alpha \mathbf{I} + \mathbf{Q}_i) \mathbf{w}_i} , \tag{13}$$

where the optimum is given as the solution of a generalized eigenvalue problem. The parameter $\alpha \ge 0$ (or rather its inverse) can be interpreted as a Lagrange multiplier of (12) and determines the trade-off between interference suppression and a low total transmission power [7].

Finally, we mention traditional beamforming,

$$\mathbf{w}_i = \arg\max \frac{\mathbf{w}_i^* \mathbf{R}_i \mathbf{w}_i}{\mathbf{w}_i^* \mathbf{w}_i} , \tag{14}$$

i.e., to use the principal eigenvector of $\mathbf{R}_i$ as the beamforming vector.

IV. Uplink Beamforming and Power Control

The problem of joint optimal power control (at the mobiles) and beamforming (at the base stations) for transmission from mobiles to base station has been considered by Rashid-Farrokhi et.al. in [20]. Here we provide additional insight into the problem by showing that it is equivalent to a convex semidefinite program.

Let $\mathbf{R}_{i,k}$ denote the channel covariance matrix from mobile i to base station k including path losses, P_i be the transmit power at mobile i and $\mathbf{w}_i$ be the receive beamformer at base station $\kappa(i)$ for the signal from user i. In the uplink, it is possible to track the instantaneous channel realization if the fading is slow enough, thus we could make the assumption that all channels are rank one. However, our formulation will hold for any channel rank.

The problem is to minimize the total transmitted power at the mobiles while it is still possible to find receive beamformers $\mathbf{w}_i$ that give a high enough SINR, $\text{SINR}_i \geq \gamma_i$, at the respective base stations, i.e.,

$$\begin{aligned} &\min \sum_{i=1}^{d} P_i \\ &\text{s.t. } \frac{\mathbf{w}_i P_i \mathbf{R}_i \mathbf{w}_i}{\sum_{n \neq i} \mathbf{w}_i^* P_n \mathbf{R}_{n,\kappa(i)} \mathbf{w}_i + \sigma^2_{\kappa(i)} \mathbf{w}_i^* \mathbf{w}_i} \geq \gamma_i \\ &\quad \|\mathbf{w}_i\| = 1 \end{aligned} \tag{15}$$

or equivalently

$$\begin{aligned} &\min \sum_{i=1}^{d} P_i \\ &\text{s.t. } \mathbf{w}_i^* \left(\gamma_i \sigma^2_{\kappa(i)} \mathbf{I} + \gamma_i \sum_{n \neq i} P_n \mathbf{R}_{n,\kappa(i)} - P_i \mathbf{R}_i \right) \mathbf{w}_i \leq 0 \\ &\quad \|\mathbf{w}_i\| = 1 \; . \end{aligned} \tag{16}$$

Just as in (6), all constraints must hold with equality at the optimum and Lemma A.1 shows (with minor modifications) that the optimal P_i are also given by the following semidefinite problem.

$$\begin{aligned} &\max \sum_{i=1}^{d} P_i \\ &\text{s.t. } \gamma_i \sigma^2_{\kappa(i)} \mathbf{I} + \gamma_i \sum_{n \neq i} P_n \mathbf{R}_{n,\kappa(i)} - P_i \mathbf{R}_i \succeq 0 \; . \end{aligned} \tag{17}$$

Once the optimal P_i have been calculated, any vector in the null space of $\mathbf{Z}_i = \gamma_i \sigma^2_{\kappa(i)} \mathbf{I} + \gamma_i \sum_{n \neq i} P_n \mathbf{R}_{n,\kappa(i)} - P_i \mathbf{R}_i$ can be used as beamformer $\mathbf{w}_i$. Note that if $\mathbf{R}_i$ has rank one, then all but one eigenvalue of $\mathbf{Z}_i$ must be strictly positive and $\mathbf{w}_i$ will be uniquely determined (up to a scaling) as was shown in [20]. However, if $\mathbf{R}_i$ has rank higher than one, then $\mathbf{w}_i$ is not necessarily unique.

Further research is needed on how this convex formulation of the problem may be exploited in the design of efficient practical algorithms for the problem.

V. Numerical Examples

We illustrate the performance of the suggested downlink algorithms in a simulated scenario with three users served by a single base station. One mobile is located at $\theta_1 = 10°$ relative array broadside and the two others at directions $\theta_{2,3} = 10° \pm \Delta$, where$\Delta$ is varied from 5°to 25°. The transmitting antenna array is linear and has $m = 8$ elements spaced half a wavelength. Each user is surrounded by a large number of local scatterers corresponding to a spread angle of $\sigma_\theta = 2°$, as seen from the base station. Thus, the channel covariance matrix is well approximated by [1], [23],

$$[\mathbf{R}(\theta, \sigma_\theta)]_{kl} = e^{j\pi(k-l)\sin\theta} e^{-\frac{(\pi(k-l)\sigma_\theta \cos\theta)^2}{2}} \; .$$

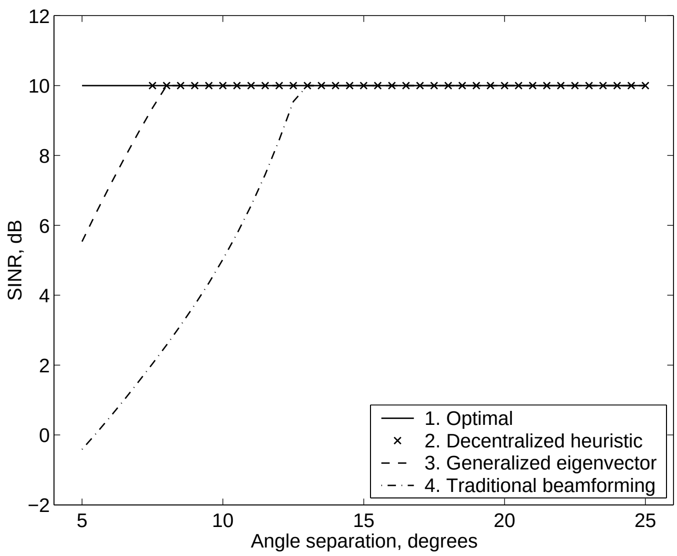

Fig. 1. Received signal to interference ratio (same for all users).

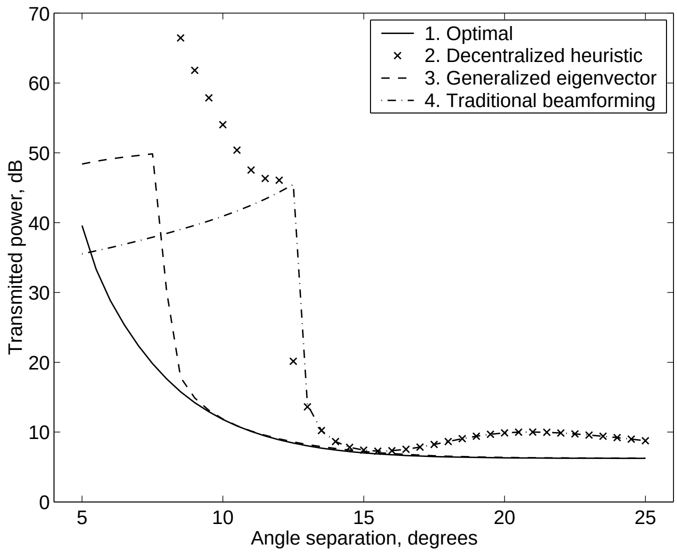

Fig. 2. Total transmitted power, relative the noise level at each receiver.

We compare the following four transmit strategies.

1. The joint optimal design according to Section III.III-A.
2. The decentralized heuristic design given by (12).
3. The generalized eigenvector beamformer given by (13).
4. Traditional beamforming given by (14).

In all cases, the beamformers are scaled, if possible, such that $\text{SINR}_i = \gamma_i$ for all users. In the cases where no such scaling is possible, the scaling is chosen to maximize the worst SINR among the users, thus all users receive the same SINR in all the examples. The semidefinite problems were solved using a primal-dual interior point method [21]. Note that the algorithm in [19] would have provided exactly the same solution.

The SINR threshold was set to $\gamma_i = 10\text{dB}$ and in (12), the thresholds were set to $\mu_i \sigma_i^2/\xi_i = 10$dB. In (13), the parameter α was arbitrarily set to

$$\alpha = \frac{0.1}{m} \operatorname{Tr}\Big[\sum_{n \neq k}^{d} \mathbf{R}_n\Big] , \tag{18}$$

10% of the gain level of the interfered channels.

Figure 1 shows the received SINR level for the users. The total transmitted power needed to achieve this level for all users is shown in Figure 2.

Because of the scaling used for the solutions, Figure 1 mainly shows when it is possible to achieve the desired SINR at all receivers using the different strategies.

As could be expected, the traditional beamformer performs worst, since it does not take the interfered users into account. However, when the users are sufficiently separated such that the traditional beamformer does give a feasible solution, then it coincides with the heuristic solution in (12), which thus can be seen as a generalization of the traditional beamforming. The performance of the generalized eigenvector solution in (13) depends on the choice of α, but in this example it performs almost as well as the optimal solution for the cases where it gives a feasible solution. In the difficult scenarios where the users are very closely separated, the cost, in terms of total transmitted power, is very high in order to obtain a feasible solution.

A thorough investigation of the properties of the constrained and robust beamformers introduced in Sections III.III-B and III.III-C falls beyond the scope of this paper.

VI. Conclusions

We have proposed strategies and algorithms for the design of downlink beamformers that keep the total transmitted power at a minimum, still maintaining a certain level of quality. The technique can be applied both on systems with inter-cell channel reuse as well as systems with at most one co-channel user per cell.

The quality constraints are given in terms of the average signal to interference plus noise ratio received at each mobile. We have shown how the resulting optimization problem can be efficiently solved using standard tools for semidefinite optimization. This is a surprising result, since the original problem is non-linear and non-convex. We have also shown a similar result for the problem of joint power control and beamforming for the uplink.

From practical experiments, the algorithm using convex optimization typically converges in about 10 iterations, whereas for the previously published algorithm [19], although the convergence is geometrical, a large number of iterations may be required if γ_i is chosen such that the problem is nearly infeasible. Also, the convex optimization routine will quickly detect the infeasible situations where no solution can be found. However, the main advantage is the flexibility offered in the choice of constraints. Also, robust beamforming is easily incorporated. Even though the method will not always produce a normal time-invariant beamformer when extra constraints are added, the solution does still provide a useful benchmark for e.g. capacity studies.

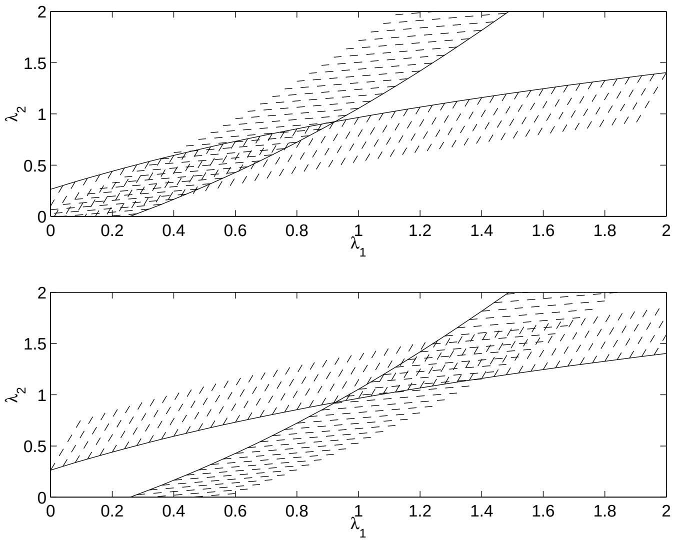

Fig. 3. Example of the feasible regions of the two problems (19), top, and (20), bottom.

Appendix
Proof of Theorem III.1

The original problem has a global minimum since the cost function is continuous and the constraint set can be limited to a compact set. We wish to show that this optimal solution is also optimal for the relaxed problem. In [19] a "virtual uplink problem" is introduced in an iterative solution for the downlink problem. Here, we will show algebraically that the downlink and the virtual uplink problems are equivalent. Then the following lemma is used to show that the virtual uplink problem is equivalent to the dual of (6), hence the duality gap is zero which also shows that the semidefinite relaxation gives the same optimal solution.

Lemma A.1: The following two problems have the same unique optimal solution $\boldsymbol{\lambda} = [\lambda_1, \lambda_2, \ldots, \lambda_d]$.

$$\begin{aligned} &\min_{\lambda_i} \sum_{i=1}^{d} \lambda_i \gamma_i \sigma_i^2 \\ &\text{s.t. } \mathbf{u}_i^* (\mathbf{I} - \lambda_i \mathbf{R}_{i\kappa(i)} + \sum_{n \neq i} \lambda_n \gamma_n \mathbf{R}_{n\kappa(i)}) \mathbf{u}_i = 0, \\ &\quad \|\mathbf{u}_i\| = 1 \\ &\quad \lambda_i \geq 0, \ i = 1, \ldots, d \end{aligned} \tag{19}$$

$$\begin{aligned} &\max_{\lambda_i} \sum_{i=1}^{d} \lambda_i \gamma_i \sigma_i^2 \\ &\text{s.t. } \mathbf{I} - \lambda_i \mathbf{R}_{i\kappa(i)} + \sum_{n \neq i} \lambda_n \gamma_n \mathbf{R}_{n\kappa(i)} \succeq 0, \ i = 1, \ldots, d \end{aligned} \tag{20}$$

Proof: First note that both problems have optimal solutions in the set given by the following constraints.

$$\begin{aligned}
&\mathbf{I}-\lambda_i\mathbf{R}_{i\kappa(i)}+\sum_{n\neq i}\lambda_n\gamma_n\mathbf{R}_{n\kappa(i)}\succeq 0,\ i=1,\ldots,d\\
&\mathbf{I}-\lambda_i\mathbf{R}_{i\kappa(i)}+\sum_{n\neq i}\lambda_n\gamma_n\mathbf{R}_{n\kappa(i)}\not\succ 0,\ i=1,\ldots,d\\
&\lambda_i>0.
\end{aligned}\tag{21}$$

If we for example assume that the factor within the parenthesis of constraint number i in (19) has both positive and negative eigenvalues at the optimum, then λ_i can be decreased without making any of the constraints strictly definite which gives a contradiction. A similar argument shows the result for (20). Assume now that we can find two different vectors $\boldsymbol{\lambda}^1\neq\boldsymbol{\lambda}^2$ in the set given by (21). Since no λ_i is zero, we can always find an i and a constant α such that $\lambda_i^1=\alpha\lambda_i^2$ and $\lambda_n^1\leq\alpha\lambda_n^2$ for all $n\neq i$. Assume without loss of generality that $\alpha>1$. Then constraint i gives

$$\begin{aligned}
\alpha\lambda_i^2\mathbf{R}_{i\kappa(i)}&=\lambda_i^1\mathbf{R}_{i\kappa(i)}\preceq\mathbf{I}+\sum_{n\neq i}\lambda_n^1\gamma_n\mathbf{R}_{n\kappa(i)}\\
&\preceq\mathbf{I}+\sum_{n\neq i}\alpha\lambda_n^2\gamma_n\mathbf{R}_{n\kappa(i)}\prec\alpha(\mathbf{I}+\sum_{n\neq i}\lambda_n^2\gamma_n\mathbf{R}_{n\kappa(i)})
\end{aligned}$$

which contradicts that $\boldsymbol{\lambda}^2$ is in the set. Thus (21) defines a single unique point which is the optimal solution to both (19) and (20). The result is illustrated in Figure 3 for a two-user example. □

Introduce separate variables for the power $P_i\geq 0$ and the normalized beamformers $\mathbf{u}_i;\|\mathbf{u}_i\|=1$ such that $\mathbf{w}_i=P_i\mathbf{u}_i$. Also, introduce the vectors $\mathbf{P}=[P_1,\ldots,P_d]^T$, $\mathbf{N}=[\gamma_1\sigma_1^2,\ldots,\gamma_d\sigma_d^2]^T$, $\boldsymbol{\omega}=[\|\mathbf{u}_1\|^2,\ldots,\|\mathbf{u}_d\|^2]^T$ and the matrix $\mathbf{F}$ where

$$[\mathbf{F}]_{kl}=\begin{cases}\mathbf{u}_k^*\mathbf{R}_k\mathbf{u}_k & k=l\\ -\gamma_k\mathbf{u}_l^*\mathbf{R}_{k\kappa(l)}\mathbf{u}_l & k\neq l\ .\end{cases}$$

Since all constraints in (6) are fulfilled with equality at the optimum, the problem can be written

$$\begin{aligned}
&\min\ \mathbf{P}^T\boldsymbol{\omega}\\
&\text{s.t.}\ \mathbf{F}\mathbf{P}=\mathbf{N}\\
&\qquad\|\mathbf{u}_i\|=1\\
&\qquad P_i\geq 0\ .
\end{aligned}\tag{22}$$

$\mathbf{F}$ could be written $\mathbf{F}=\mathbf{D}(\mathbf{I}-\mathbf{G})$, where $\mathbf{D}$ is a diagonal matrix with the same diagonal as $\mathbf{F}$ and $\mathbf{G}$ has only non-negative elements. In [19], it is shown that if the problem is feasible, then at the optimum, the spectral radius of $\mathbf{G}$ is <1 and the Frobenius Perron theory for matrices with non-negative elements [8] show that $\mathbf{F}$ is invertible and $\mathbf{F}^{-1}$ has only non-negative elements, thus the constraints $P_i\geq 0$ are fulfilled implicitly and the problem can be rewritten as

$$\begin{aligned}
&\min\ \mathbf{N}^T\mathbf{F}^{-T}\boldsymbol{\omega}\\
&\text{s.t.}\ \|\mathbf{u}_i\|=1\ .
\end{aligned}\tag{23}$$

Introducing the vector $\boldsymbol{\rho}=\mathbf{F}^{-T}\boldsymbol{\omega}$, which implicitly will be all non-negative, we finally arrive at

$$\begin{aligned}
&\min\ \mathbf{N}^T\boldsymbol{\rho}\\
&\text{s.t.}\ \mathbf{F}^T\boldsymbol{\rho}=\boldsymbol{\omega}\\
&\qquad\|\mathbf{u}_i\|=1\\
&\qquad\rho_i\geq 0\ ,
\end{aligned}\tag{24}$$

i.e.,

$$\begin{aligned}
&\min_{\rho_i}\sum_{i=1}^d\rho_i\gamma_i\sigma_i^2\\
&\text{s.t.}\ \mathbf{u}_i^*(\mathbf{I}-\rho_i\mathbf{R}_{i\kappa(i)}+\sum_{n\neq i}\rho_n\gamma_n\mathbf{R}_{n\kappa(i)})\mathbf{u}_i=0,\\
&\qquad\|\mathbf{u}_i\|=1\\
&\qquad\rho_i\geq 0,\ i=1,\ldots,d\ .
\end{aligned}\tag{25}$$

Next, note that the Lagrangian dual of (6), which is also the Lagrangian dual of the relaxation (8), is given by

$$\begin{aligned}
&\max_{\lambda_i}\sum_{i=1}^d\lambda_i\gamma_i\sigma_i^2\\
&\text{s.t.}\ \mathbf{I}-\lambda_i\mathbf{R}_{i\kappa(i)}+\sum_{n\neq i}\lambda_n\gamma_n\mathbf{R}_{n\kappa(i)}\succeq 0,\ i=1,\ldots,d\ .
\end{aligned}\tag{26}$$

According to Lemma A.1 the dual has the same optimal solution as the virtual uplink problem (25) and since the dual gives a lower bound to the relaxation (8) which in its turn is a lower bound of (6), they must all be the same. Thus, if the problem is feasible, any optimal solution to the original problem will also be in the set of optimal solutions to the relaxation. Conversely, if the relaxation is infeasible, then clearly also (6) is infeasible, which concludes the proof. Note that the strict duality between the relaxation and the dual also follows from Slater's constraint qualification [6] which is fulfilled if the problem is feasible.

Finally, we can note that any $\mathbf{u}_i$ in the null space of $\mathbf{Z}_i=\mathbf{I}-\rho_i\mathbf{R}_{i\kappa(i)}+\sum_{n\neq i}\rho_n\gamma_n\mathbf{R}_{n\kappa(i)}$ will be an optimal solution of (25). The complementarity conditions between the relaxation and the dual shows that at the optimum, $\mathrm{Tr}[\mathbf{W}_i\mathbf{Z}_i]=0$, i.e., $\mathbf{W}_i\mathbf{Z}_i=0$ since both matrices are positive semidefinite. Thus, if our semidefinite problem gives a high rank solution, any $\mathbf{u}_i\in\mathrm{span}[\mathbf{W}_i]$ will solve the virtual uplink problem and the corresponding scaling factors P_i are easily calculated from the system of linear equations $\mathbf{F}\mathbf{P}=\mathbf{N}$.

Acknowledgment

The authors would like to express their gratitude to professor Stephen Boyd at the Information Systems Laboratory, Stanford University, to associate professor Anders Forsgren and his colleagues at the division of optimization and systems theory and to associate professor Seong-Lyun Kim at the Radio Communication Systems Group, both at the Royal Institute of Technology, Stockholm, for their valuable ideas and comments.

References

[1] F. Adachi, M.T. Feeney, A.G. Williamson, and J.D. Parsons. Cross correlation between the envelopes of 900 MHz signals received at a mobile radio base station site. *IEE Proceedings, Pt. F*, 133(6):506–512, October 1986.

[2] Kurt Anstreicher and Henry Wolkowicz. On Lagrangian relaxation of quadratic matrix constraints. Technical Report CORR 98-24, Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario, Canada, 1998. To appear in SIAM Journal on Matrix Analysis and Applications.

[3] T. Asté, P. Forster, L. Féty, and S. Mayrargue. Downlink beamforming avoiding DOA estimation for cellular mobile communications. In *Proc. IEEE ICASSP 98*, volume 6, pages 3313–3316, 1997.

[4] Mats Bengtsson and Björn Ottersten. On approximating a spatially scattered source with two point sources. In *Proc. NORSIG'98*, pages 45–48, June 1998.

[5] Mats Bengtsson and Björn Ottersten. Downlink beamformer design using semidefinite optimization. In *Proceedings of RadioVetenskap och Kommunikation (RVK)*, pages 289–293, Karlskrona, Sweden, June 1999.

[6] Dimitri P. Bertsekas. *Nonlinear Programming*. Athena Scientific, Belmont, Massachusetts, 1995.

[7] Henry Cox, Robert M. Zeskind, and Mark M. Owen. Robust adaptive beamforming. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 35(10):1365–1375, October 1987.

[8] F. R. Gantmacher. *The theory of matrices.*, volume II. Chelsea Publishing Company, 1959.

[9] Lal C. Godara. Applications of antenna arrays to mobile communications. I. Performance improvement, feasibility, and system considerations. *Proceedings of IEEE*, 85(7):1031–1060, July 1997.

[10] Jason Goldberg and Javier R. Fonollosa. Downlink beamforming for cellular mobile communications. In *Proceedings of VTC'97*, volume 2, pages 632–636, 1997.

[11] Hervé Lebret and Stephen Boyd. Antenna array pattern synthesis via convex optimization. *IEEE Transactions on Signal Processing*, 45(3):526–532, March 1997.

[12] Miguel Sousa Lobo, Lieven Vandenberghe, Stephen Boyd, and Hervé Lebret. Applications of second-order cone programming. *Linear Algebra Appl.*, 284(1-3):193–228, 1998.

[13] Hussein K. Mecklai and Rick S. Blum. Transmit antenna diversity for wireless communications. In *Proceedings IEEE International Conference on Communications 95*, volume 3, pages 1500 – 1504, 1995.

[14] Guiseppe Montalbano and Dirk T.M. Slock. Spatio-temporal array processing for matched filter bound optimization in SDMA downlink transmission. In *URSI Internat. Symp. Signals, Systems and Electronics*, pages 416 – 421, 1998.

[15] Robert A. Monzingo and Thomas W. Miller. *Introduction to Adaptive Arrays*. Wiley, 1980.

[16] Aradhana Narula, Michael J. Lopez, Mitchell D. Trott, and Gregory W. Wornell. Efficient use of side information in multiple-antenna data transmission over fading channels. *IEEE Journal on Selected Areas in Communications*, 16(8):1423–1436, October 1998.

[17] Yurii Nesterov and Arkadii Nemirovskii. *Interior-point polynomial algorithms in convex programming*. SIAM Publications, Philadelphia, PA, 1994.

[18] Farrokh Rashid-Farrokhi, K.J. Ray Liu, and Leandros Tassiulas. Transmit beamforming for cellular wireless communications. In *Proc. Conf. Info. Sciences and Systems (CISS-97)*, pages 92–97, Johns Hopkins, Baltimore, March 1997.

[19] Farrokh Rashid-Farrokhi, K.J. Ray Liu, and Leandros Tassiulas. Transmit beamforming and power control for cellular wireless systems. *IEEE Journal on Selected Areas in Communications*, 16(8):1437–1450, October 1998.

[20] Farrokh Rashid-Farrokhi, Leandros Tassiulas, and K.J. Ray Liu. Joint optimal power control and beamforming in wireless networks using antenna arrays. *IEEE Transactions on Communications*, 46(10):1313–1324, October 1998.

[21] Jos F. Sturm. *Using SeDuMi 1.0x, a MATLAB Toolbox for Optimization over Symmetric Cones*. Department of Quantitative Economics, Maastricht University, The Netherlands, `http://www.unimaas.nl/~sturm/software/sedumi.html`, 1999.

[22] Vahid Tarokh, Nambi Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Transactions on Information Theory*, 44(2):744 – 765, March 1998.

[23] Tõnu Trump and Björn Ottersten. Estimation of nominal direction of arrival and angular spread using an array of sensors. *Signal Processing*, 50(1-2):57–69, April 1996.

[24] Lieven Vandenberghe and Stephen Boyd. Semidefinite programming. *SIAM Review*, 38(1):49–95, March 1996.

[25] Eugene Visotsky and Upamanyu Madhow. Optimum beamforming using transmit antenna arrays. In *Proceedings of VTC'99*, volume 1, pages 851–856, Houston, Texas, USA, May 1999. IEEE.

[26] Gregory W. Wornell and Mitchell D. Trott. Efficient signal processing techniques for exploiting transmit antenna diversity on fading channels. *IEEE Transactions on Signal Processing*, 45(1):191–204, January 1997.

[27] Per Zetterberg. *Mobile Cellular Communications with Base Station Antenna Arrays: Spectrum Efficiency, Algorithms and Propagation Models*. PhD thesis, Royal Institute of Technology, Stockholm, Sweden, June 1997.

[28] Per Zetterberg and Björn Ottersten. The spectrum efficiency of a base station antenna array for spatially selective transmission. *IEEE Transactions on Vehicular Technology*, 44:651–660, August 1995.