ADAPTIVE BEAMFORMING AND SPATIAL MULTIPLEXING OF UNICAST AND MULTICAST SERVICES

Yuri C. B. Silva* and Anja Klein
Darmstadt University of Technology, Merckstr. 25, 64283 Darmstadt, Germany
Email: y.silva@nt.tu-darmstadt.de, a.klein@nt.tu-darmstadt.de

Abstract
This paper evaluates and compares different adaptive antenna techniques applied in the context of multicast services and presents a methodology for performing the spatial multiplexing of both unicast and multicast users. It is seen that adaptive beamforming is able to provide good results even for large groups of multicast users and that the presence of line-of-sight is beneficial to the algorithms which focus on the performance of the worst user. Additionally, grouping strategies that allow the allocation of the same resources to unicast and multicast users, and which make use of the proposed spatial multiplexing procedure, are shown to be more efficient than allocating separate resources to unicast and multicast users.

I. INTRODUCTION
In the context of next-generation wireless systems, it is expected that services targeted at mass content distribution become widely popular, especially considering the 3GPP standardization activities for their implementation within GSM/EDGE and WCDMA networks [1]. Examples of such services are audio/video streaming, mobile TV, messaging, news clips, localized services, download, among others. Their common characteristic is that the same information has to be transmitted to a group (multicast) or to all users (broadcast) within a certain coverage area.

The implementation of such services in mobile cellular networks raises some issues concerning their feasibility and efficiency within the different layers of the transmission chain [1]. The scope of this article lies on the radio link between the mobile and base stations, for which one of the main issues concerns the optimization of the use of the radio resources for multicast services.

This paper focuses on the application of adaptive antenna arrays to multicast services and also on the spatial multiplexing of groups containing both unicast and multicast users. First it is assumed that only multicast users share the same resource, for which adequate adaptive beamforming techniques are presented and evaluated. Then, it is considered that unicast users are allowed to share the same resource with the multicast users. A spatial multiplexing procedure is therefore proposed for allowing the coexistence of unicast and multicast. It consists of the following three steps: block diagonalization, multicast beamforming, and power loading.

The paper is organized as follows. In section II, different adaptive beamforming techniques are formulated for multicast services. Section III presents a strategy for the spatial multiplexing of unicast and multicast users. The performance evaluation results are presented in section IV. Finally, section V draws some conclusions.

II. ADAPTIVE BEAMFORMING FOR MULTICAST
A multi-user multi-carrier system is considered, which assumes flat-fading per sub-carrier and negligible inter-symbol interference (ISI), so that the data symbols can be treated individually. The base station has an antenna array composed of \( M \) elements and serves a group of \( N \) single-antenna users. For the unicast case, considering a vector \( \mathbf{d}_{N \times 1} \) with \( N \) data symbols (each addressed to a different user), which are modulated by a matrix \( \mathbf{M}_{M \times N} \), transmitted over the radio channel \( \mathbf{H}_{N \times M} \), subject to additive white Gaussian noise \( \mathbf{n}_{N \times 1} \), and demodulated by a matrix \( \mathbf{D}_{N \times N} \), the \( N \) downlink estimates \( \hat{\mathbf{d}}_{N \times 1} \) of the \( N \) transmitted symbols \( \mathbf{d} \) may be written as

\[
\hat{\mathbf{d}} = \mathbf{DHMd} + \mathbf{Dn}.
\]

The multicast scenario can then be seen as a particular case of a MIMO multiuser system [2], for which all users expect the same symbols, i.e., \( \mathbf{d} = s \mathbf{1} \), where \( \mathbf{1}_{N \times 1} \) is a vector of ones and \( s \) is the data symbol. Equation (1) may then be rewritten as

\[
\hat{\mathbf{d}} = \mathbf{DHws} + \mathbf{Dn},
\]

where \( \mathbf{w}_{M \times 1} = \mathbf{M} \mathbf{1} \) is the resulting weight vector, which is the sum of the weight vectors of the individual users contained within \( \mathbf{M} \).

The objective of the algorithms within this section is to perform multicast transmit processing, i.e., to properly determine the transmit weight vector \( \mathbf{w} \) in (2). However, since the number of users within a multicast group may well exceed the number of transmit antennas, it is not possible in such cases to guarantee that the received symbols will be in-phase with the actual symbols through pure transmit processing. The following diagonal demodulation matrix is therefore considered

\[
\mathbf{D} = \text{diag}(|\mathbf{H}_1|, \ldots, |\mathbf{H}_N|)\mathbf{H}^H \text{diag}(\mathbf{H})^{-1},
\]

where \( \mathbf{H}_{1 \times M} \) corresponds to the channel of user \( i \), i.e., the \( i^{th} \) row of matrix \( \mathbf{H} \). \( \text{diag}(\cdot) \) is a diagonal matrix with the arguments on the diagonal and \( (\cdot)^H \) denotes the conjugate transpose of a matrix.
In the following subsections, the herein investigated multicast transmit processing algorithms are formulated. Since the multicast users expect the same symbols, algorithms which try to suppress intracell interference, such as zero-forcing and Tomlinson-Harashima precoding, are not considered.

The matched filter is presented first, being followed by the multicast-specific algorithms, which aim at the maximization of the lowest signal-to-noise ratio (SNR) among the multicast connections [3-5], and finally, the User Selective Matched Filter (USMF) is proposed.

A. Matched filter

The matched filter optimization for a single user scenario consists of finding the weight vector which maximizes the SNR perceived at the receiver. In the case of multicast, equivalently, it can be expressed as the maximization of the sum, or average, SNR perceived by the users within the multicast group [6]. The optimization problem may thus be written as

$$w_{opt} = \arg \max_w \frac{E\{|\mathbf{H}w_s|^2\}}{E\{|\mathbf{n}|^2\}}$$

subject to: $||w||^2 \leq E_t$,  

where $w_{opt}$ is the optimal weight vector, $E_t$ is the available transmit energy, $E\{\cdot\}$ is the expectation operator, and $|| \cdot ||$ is the Euclidean norm of a vector. This optimization leads to an eigenvalue problem, with solution

$$w_{opt} = \beta \cdot (\text{principal eigenvector of } \mathbf{H}^H \mathbf{H}),$$

$$\beta = \sqrt{E_t/\sigma_n^2},$$

where $\sigma_n^2$ is the average symbol energy [6].

B. Max-min algorithms

The quality perceived by the users within a multicast group may vary significantly, depending on their radio channel conditions. Fairness among the users could therefore be introduced by the following optimization procedure, which tries to maximize the minimum SNR within the group:

$$w_{opt} = \arg \max_w \min \{SNR_i\}$$

with $SNR_i = |\mathbf{H}_i w_s|^2/\sigma_n^2$, $i = 1, \ldots, N$

subject to: $||w||^2 \leq E_t$,  

where $\sigma_n^2$ is the noise power.

This optimization problem is a quadratically constrained quadratic programming problem and does not have a closed-form solution. In [3] the problem is solved using sequential quadratic programming, while other articles [4, 5] have presented different approaches for simplifying the problem and finding more computationally efficient solutions. In [4], for example, the problem is relaxed by removing one of the constraints of an equivalent optimization problem, which can then be solved efficiently by semidefinite programming methods.

C. User selective matched filter

In this paper we propose a heuristic algorithm called User Selective Matched Filter (USMF), which does not claim to provide the optimum for (6), but which tries to improve the performance of the matched filter in a multicast scenario.

If it were assumed that there is a point-to-point connection for each user $i$, the ideal solution in the sense of maximizing the SNR would be to employ a transmit matched filter, i.e., $w = \mathbf{H}_i^H$. The idea of USMF is to stack these individual weight vectors within a matrix ($\mathbf{H}^H$), but disregarding the weight vectors within it that do not positively contribute to the goal of maximizing the lowest SNR. The algorithm may be written as

$$w_{USMF} = \beta \cdot \mathbf{H}^H \mathbf{P}_1,$$

$$\beta = \sqrt{(E_t/\sigma_n^2) \cdot tr(\mathbf{P}^T \mathbf{H}^H \mathbf{P} \mathbf{1}^T)},$$

where $\mathbf{P}_{N \times N}$ is a non-zero diagonal matrix, with elements $p_{ii} \in \{0, 1\}$, for $i = 1, \ldots, N$. Since there are $N$ users, and the diagonal elements of $\mathbf{P}$ are restricted to binary values, there exists a total of $2^N - 1$ possibilities.

For small group sizes, all possible $\mathbf{P}$ matrices can be evaluated, from which the one providing the highest minimum SNR can be chosen. However, complexity grows exponentially with an increasing number of users. An alternative for making it computationally efficient would be to evaluate only a limited number of possibilities, which could for example be selected through randomization.

III. SPATIAL MULTIPLEXING

The previous section has dealt with transmit processing techniques adequate for the provision of multicast services. In practice, however, such services will coexist with traditional unicast point-to-point connections. In order to support both services efficiently, assuming that multiple antennas are available at the base station, spatial multiplexing techniques may be employed to improve system capacity.

Space division multiple access (SDMA) techniques have already been extensively studied for multi-antenna unicast scenarios [7, 8]. The intra-cell interference that arises from the simultaneous use of the radio resources by multiple users can be mitigated through algorithms such as zero-forcing, Tomlinson-Harashima precoding, block diagonalization, among others [9-11].

When both unicast and multicast users build an SDMA group, such interference suppression techniques do also apply. The difference is that there is no need to suppress the interference among the multicast users, which expect the same data stream, i.e., only the interference between unicast and multicast, and interference among the unicast users has to be mitigated. Such constraints lead to a block diagonal structure, similar to that presented in [11] for MIMO unicast users, but with a large block composed of the multicast users and small individual unicast blocks.
This block diagonal approach for unicast/multicast has been first suggested in [12]. Such structure allows that the blocks be individually processed, i.e., the beamforming algorithms presented in section II may be directly applied to the multicast group.

For that purpose we propose that the spatial multiplexing of a unicast/multicast SDMA group be divided into three steps: diagonalization, beamforming, and power loading. They are represented, respectively, by matrices $N_{M \times N}$, $B_{N \times N}$, and $\Gamma_{N \times N}$, with $N$ denoting the total number of users ($N_{mc}$ multicast plus $N_{uc}$ unicast users). The system equation in (1), when defining $M = \beta N B \Gamma$, becomes

$$\hat{d} = D H(\beta N B \Gamma) d + Dn,$$

where $\hat{d}$, $d$, $D$, and $n$ preserve the dimensions defined in section II. The normalization factor $\beta$ is introduced in order to satisfy the transmit energy constraint, $||\beta N B \Gamma d||^2 = E_{tr}$, and is given by

$$\beta = \sqrt{E_{tr}/tr \{ (N B \Gamma)^H (N B \Gamma) R_d \},}$$

where $R_d$ is the covariance matrix of the signal vector $d$.

The procedure for determining the spatial multiplexing matrices, as well as the discussion on possible grouping strategies for unicast/multicast, are approached in the following subsections. Note that the existence of only one multicast group is assumed, but the procedure can be directly extended to any given number of groups.

A. Diagonalization

The block diagonalization algorithm presented in [11] can be applied to the combination of unicast and multicast by considering one multicast block of size $B_j = N_{mc}$ and $N_{uc}$ unicast blocks of size $B_j = 1$ (single-antenna terminals), with $B_j$ denoting the size of block $j$. The total number $N_b$ of blocks is therefore equal to $N_{uc} + 1$ and $\sum_{j=1}^{N_b} B_j = N$.

The matrix $H_{j} \times B_j \times M$ corresponds to the channel matrix of block $j$, while $\tilde{H}_{j} \times (N-B_j) \times M$ denotes the composed channel matrix of all other blocks:

$$\tilde{H}_{j} = [H_{j}^T \ldots H_{j-1}^T \ldots H_{j+1}^T \ldots H_{N_b}^T]^T. \quad (10)$$

According to the block diagonalization algorithm, the diagonalization matrix $N$ can be written as

$$N = [(\tilde{V}_{j}^{(0)} \cdot V_{j}^{(1)}) \ldots (\tilde{V}_{N_b}^{(0)} \cdot V_{N_b}^{(1)})], \quad (11)$$

where $\tilde{V}_{j}^{(0)}$ corresponds to the null space of $\tilde{H}_{j}$ and $V_{j}^{(1)}$ is the signal space of $H_{j}^{(0)}$.

The equivalent channel matrix after the diagonalization, if we consider as an example a system with $N_{mc} = N_{uc} = 2$ users, should have the following structure, with $x$ representing non-zero matrix entries,

$$HN = \begin{bmatrix} x & x & 0 & 0 \\ x & x & 0 & 0 \\ 0 & 0 & x & 0 \\ 0 & 0 & 0 & x \end{bmatrix}. \quad (12)$$

B. Beamforming

After the diagonalization is performed and the users are separated in blocks, additional transmit processing may be done for the $N_{mc} \times N_{mc}$ multicast block $(HN)_{mc}$. The multicast beamforming algorithms presented in section II may be applied to improve the group performance, resulting in an $N_{mc} \times N_{mc}$ beamforming matrix $B_{mc}$. The complete matrix is block diagonal and can be written as

$$B = \begin{bmatrix} B_{mc} & 0^T \\ 0 & I_{N_{uc}} \end{bmatrix}, \quad (13)$$

where $I_{N_{uc}}$ denotes an $N_{uc} \times N_{uc}$ identity matrix, and $0$ is an $N_{uc} \times N_{mc}$ matrix with zeros.

C. Power loading

The last step of the spatial multiplexing consists of the distribution of the available transmit power among the different blocks. The allocation strategies known for the unicast case, e.g., waterfilling, equal power load, and equal received power [2], can also be applied to the unicast/multicast scenario with some minor modifications.

The application of waterfilling results in the maximization of the sum capacity, but on the other hand it may lead to users with largely varied quality profiles. The allocation of the same amount of power to the users is the simplest procedure, but may also lead to an uneven quality distribution, due to the different channel attenuation associated to each user. A fair scheme for the power distribution among the blocks is here proposed, which tries to guarantee that the unicast users will receive the same signal power as the worst user within the multicast group.

The power loading matrix $\Gamma$ is a diagonal and real matrix. The elements associated to the same diagonalization block are said to be identical, i.e., the division of the power among the users of the multicast group is assumed to have already been done by the multicast beamforming algorithm. The $\Gamma$ matrix can be expressed as

$$\Gamma = \begin{bmatrix} \gamma_1 I_{B_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \gamma_{N_b} I_{B_{N_b}} \end{bmatrix}, \quad (14)$$

with each $\gamma_j$ given by

$$\gamma_j = \frac{1}{\sum_{k=1}^{N_b} B_k (X_j/X_k)^2}, \quad (15)$$

where $(HN B d)_j$ is a $B_j \times 1$ vector, corresponding to the signal that is estimated to be received by each user within block $j$, and $(HN B d)_j,n$ is the signal of the $n^{th}$ user within the same block $j$. 

The 17th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC’06)
where $H$ of deterministic-to-scattered power, of zero mean circularly symmetric complex Gaussian unicast and multicast users, similar to a unicast-only case, is

The size of a spatial multiplexing group containing both unicast and multicast users, similar to a unicast-only case, is upper-limited by the number of transmit antennas available at the base station. If more users need to be served, then other multiplexing dimensions (time or frequency) have to be taken into account.

A simple grouping approach consists of separating the users according to their type of service, i.e., unicast and multicast users are allocated to different time or frequency resources. This would mean that traditional unicast SDMA and multicast beamforming could be employed separately on their respective resources.

However, an allocation scheme which allows both unicast and multicast users to share the same resources might be more efficient than simply isolating both services. In order to implement such a scheme, the previously presented multiplexing procedure is therefore required.

In order to have an upper bound of the performance gains that an efficient grouping might provide, it is here considered that, among all possible groupings, the one providing the highest minimum user capacity is selected. The simulation results presented next consider an exhaustive group search, but other more computationally efficient schemes [8] could be employed instead.

IV. PERFORMANCE EVALUATION

The simulation scenario consists of a single cell equipped with a four-element uniform linear antenna array, single antenna mobile terminals, and QPSK modulation. The implemented channel model regards both line-of-sight (LOS) and non-line-of-sight (NLOS) components [2], and can be written as

$$H = \sqrt{K/(1+K)} \mathbf{H} + \sqrt{1/(1+K)} \mathbf{H}_w,$$  \hspace{1cm} (16)

where $K$ is the Ricean factor which determines the ratio of deterministic-to-scattered power, $\mathbf{H}_w \in \mathbb{C}^{N \times M}$ is composed of zero mean circularly symmetric complex Gaussian random variables with unit variance, and $\mathbf{H}$ models the LOS component. Note that the effects of path-loss and log-normal fading are assumed to be compensated by power control.

Initially, the multicast-only case is considered. The beamforming algorithms are implemented according to section II. The "Max-min opt." algorithm refers to the optimization problem in (6) solved through numerical optimization, and the LOS and NLOS scenarios correspond to $K \rightarrow \infty$ and $K = 0$ in (16), respectively.

Fig. 1 shows the average bit error rate (BER) performance of the different algorithms from section II for a multicast group composed of four users in NLOS and LOS scenarios, respectively. The BER is depicted as a function of the $E_s/N_0$, which represents the ratio of the symbol energy to the spectral noise density. It can be seen that the solution of the max-min problem in (6) presents the lowest bit error rates, being followed by the USMF algorithm, which for the LOS scenario requires approximately an extra 2.5dB in order to provide a BER of $5 \cdot 10^{-3}$.

When we compare the results obtained for a rich scattering scenario to those obtained for a purely line-of-sight situation, it becomes clear that the channel profile has a considerable impact on the performance of the algorithms. The USMF gets much closer to the optimal solution in the presence of LOS. The increased spatial correlation of this scenario has a positive effect on USMF, which can be explained due to the fact that it increases the probability that the rows of $\mathbf{H}^H$ be correlated, resulting in more zero entries within $\mathbf{P}$ in (7), which brings it closer to the single-user beamforming case.

The impact that the multicast group size has on the performance of the algorithms can be seen in Fig. 2, for both the NLOS and LOS scenarios. An input $E_s/N_0$ of 10dB is assumed and the results are presented in terms of the 10th percentile of the cumulative distribution of the minimum SNR within the multicast group. This indicates a 90% probability that the SNR perceived by the worst user within the group is higher than the given value.

For all algorithms it can be seen that the more users there are within a group, the lower the SNR that can be
guaranteed. Up to a certain number of users (roughly 4-5 for most algorithms) the descent is steeper, but then it tends to slowly saturate for larger group sizes.

The relative behavior among the algorithms, for both NLOS and LOS scenarios, is very similar to that verified through the BER evaluation. The performance of the USMF improves for the LOS channel, getting closer to the optimum, while the matched filter is severely degraded as the number of users increases in a channel with line-of-sight. The poor performance of the matched filter for the LOS channel is due to the fact that this scenario leads to an ill-conditioned $H^H H$ matrix, which means that more energy is concentrated on the principal eigenmode. This has a positive effect on the average SNR, but leads to a more uneven energy distribution within the group.

Fig. 3 shows a comparison between the grouping strategies discussed in section III for unicast and multicast services. The capacity of the worst user, among both unicast and multicast services, assuming Gaussian signalling and an input SNR of 10dB, is presented as a function of the number of unicast users, while the number of multicast users is fixed to 4. Users within a group are multiplexed in space, and the groups are multiplexed in time (a maximum of two groups is assumed). Note that the capacity is normalized by the number of groups in order to capture the effect of the time-multiplexing.

The joint strategy refers to the case in which users of different services may be grouped together, and the optimal grouping is found through exhaustive search, while for the separate strategy the unicast and multicast users are always in different groups. It can be seen that, as expected, the capacity decreases with an increasing number of users. The joint strategy presents better capacity results than the isolated one, especially for higher number of users. In the case of unicast, when the number of users gets closer to the number of antenna elements, it becomes more difficult to diagonalize them, and the capacity is thus significantly affected, which is not the case for multicast. Therefore it is more efficient, in terms of fairness, to group the users so that both services can be multiplexed in space.

**REFERENCES**

1. 3GPP. “Multimedia Broadcast/Multicast service (MBMS); architecture and functional description,” TS 23.246, 2005.