Comparison of Distributed Beamforming Algorithms for MIMO Interference Networks

David A. Schmidt, *Member, IEEE,* Changxin Shi, Randall A. Berry, *Senior Member, IEEE,* Michael L. Honig, *Fellow, IEEE,* and Wolfgang Utschick, *Senior Member, IEEE*

Abstract

This paper presents a comparative study of algorithms for jointly optimizing beamformers and receive filters in an interference network, where each node may have multiple antennas, each user transmits at most one data stream, and interference is treated as noise. We focus on techniques that seek good suboptimal solutions by means of iterative and distributed updates. Those include forward-backward iterative algorithms (max-Signal-to-Interference plus Noise Ratio (SINR) and interference leakage), weighted sum Mean Squared Error (MSE) algorithms, and interference pricing with incremental Signal-to-Noise Ratio (SNR) adjustments. We compare their properties in terms of convergence and information exchange requirements, and then numerically evaluate their sum rate performance averaged over random (stationary) channel realizations. The numerical results show that the max-SINR algorithm achieves the maximum Degrees of Freedom (i. e., supports the maximum number of users with near-zero interference), and exhibits better convergence behavior at high SNRs than the weighted sum MSE algorithms. However, it assumes fixed power per user, and achieves only a single point in the rate region whereas the weighted sum MSE criterion gives different points. In contrast, the incremental SNR algorithm adjusts the beam powers and deactivates users when interference alignment is infeasible.

D. A. Schmidt and W. Utschick are with the Associate Institute for Signal Processing, Technische Universität München, Munich, Germany (e-mail: dschmidt@tum.de; utschick@tum.de).

C. Shi, R. A. Berry, and M. L. Honig are with the Department of Electrical Engineering and Computer Science, Northwestern University, Evanstion, IL (e-mail: cshi@u.northwestern.edu; rberry@eecs.northwestern.edu; mh@eecs.northwestern.edu).

This work was supported in part by the NSF under grant CCF-0644344, DARPA under grant W911NF-07-1-0028, and a gift from Futurewei.

This work was presented in part at the 2009 Asilomar Conference on Signals, Systems, and Computers and the 2010 Allerton conference on Communications, Control, and Computing.

Furthermore, that algorithm can provide a slight increase in sum rate, relative to max-SINR, at the cost of additional iterations.

I. Introduction

An *interference network* consists of multiple transmitter-receiver pairs, each communicating over a wireless channel. Assuming those pairs share the same bandwidth, simultaneous transmissions cause mutual interference. Such a network includes a cellular network where interference on a given time-frequency slot originates from other cells. A challenge in such networks is then to mitigate and manage interference across nearby cells.

Allowing the nodes in an interference network to cooperate can significantly increase the overall spectral efficiency of the network [1]. "Cooperation" among nodes means that they coordinate resources including transmit power, frequencies, and beamforming coefficients when the transmitters have multiple antennas. (We do not consider message sharing across transmitters.) Specifically, we study distributed algorithms for optimizing beamformers and powers in a K-user interference network with Multi-Input/Multi-Output (MIMO) channels. Here "distributed" means that the beamformers and powers are computed (updated) locally by the associated transmitters using local channel state information. To gain some basic insights, we assume that the wireless propagation channels are narrowband and known perfectly at all wireless nodes. We also assume that the transmitters independently select their codebooks, and the receivers treat interference as noise. Although this is known to be suboptimal (i. e., does not achieve the maximum achievable sum rate), the best known coding strategies at high Signal-to-Noise Ratios (SNRs) are substantially more complicated, cf. [2]–[5]. Also, we only examine the case where each user transmits a single data stream. While this does not maximize spectral efficiency in many scenarios, this is an important case in practice, and simplifies the optimization problem substantially. (In particular, we do not consider optimization of precoder ranks, which is quite difficult in general [6], [7].)

Even with these simplifications the problem of determining the jointly optimal sets of beamformers and receivers is challenging. The sum rate objective is non-convex and has many local optima [7]–[9]. With N antennas at each node, achieving the maximum sum rate at high SNRs with number of users $K > N$ generally requires *interference alignment*; i. e., the $K - 1$ interferers (mostly) reside in a subspace with dimension at most $N-1$ [2] (provided K is not too large). In general, there can be many aligned solutions, which correspond to local optima at high SNRs, and which exhibit dramatically different performance [10], [11].

The purpose of this work is to provide a systematic comparison of distributed beamforming optimization algorithms in terms of performance, complexity, information exchange, and convergence properties (since they require iterative updates). Algorithms considered include max-Signal-to-Interference-plus-Noise-Ratio (max-SINR) [12], [13], distributed interference pricing [8], [14], weighted sum-Mean Squared Error (MSE) (developed in a conference version of this paper [15] as well as in [7], [16]), and an "incremental SNR" algorithm that attempts to track the optimum as the SNR increases incrementally. A general sum-utility criterion is considered, where each user is assigned a concave increasing utility function of SNR. In particular, the weights in the weighted sum-MSE algorithm can be updated so that minimizing weighted sum-MSE coincides with maximizing the sum utility. Related weighted sum-MSE algorithms have been presented in [16] and [7] for single and interfering MIMO broadcast channels, respectively. Here we give conditions for the convergence with single beams for an interference network and compare the performance with other algorithms.

Simulated performance (sum-rate) results are presented for interference networks with random $N \times N$ channels containing *i. i. d.* complex Gaussian elements. The results provide several insights. The following observations assume that the number of users $K = 2N - 1$, corresponding to the maximum *Degrees of Freedom (DoFs)* of the interference network. This is the maximum number of data streams (users) that can be supported without interference [10], [17], [18], so that the transmit powers are set at the maximum value.

1) The max-SINR and weighted-sum-MSE algorithms can achieve the maximum DoFs at high SNRs (i. e., suppress all interference); however, the weighted-sum-MSE algorithm generally requires many more iterations to converge.
2) The incremental SNR algorithm can provide a slight increase in sum rate, relative to max-SINR, at the cost of additional iterations. Hence the max-SINR algorithm does not generally achieve a globally optimal solution.
3) While interference pricing gives excellent performance in medium to large systems $K > 4$, it often does not achieve the maximum DoFs in a small system ($K = 3,\ N = 2$).

The weighted-MSE and interference pricing updates adjust the transmit powers jointly with the beams, whereas the max-SINR algorithm assumes fixed power for each user. (Also, max-SINR achieves only a single point in the rate region whereas weighted sum-MSE can achieve different points.) This can be an advantage if interference cannot be eliminated for an initial choice of K and N, since those algorithms along with the incremental SNR algorithm can shut off users. For this scenario we present an algorithm, which combines interference pricing for *power control* (outer loop) with the max-SINR algorithm for determining beam directions (inner loop). This effectively implements a form of admission control, although results with $K > 2N$ show that the average DoFs achieved with this method is somewhat less than the maximum (i. e., the power control loop has a tendency to turn off too many users).

There have been other algorithms proposed that we do not include in our comparison, e. g., see [19]–[25]. This is because they are either not compatible with our system assumptions (i. e., distributed algorithms with rank-one precoders), or the performance has not been observed to be better than that achieved by the algorithms considered here. Further related discussion is presented in Section IV-I.

In Sections II and III, we introduce the system model and state the sum-utility maximization problem. The algorithms considered are presented in Section IV along with a comparison of properties and information exchange requirements, and in Section V we present the simulated performance results.

II. System Model

We consider a system with K transmitter-receiver pairs (or users), where transmitter k has N_k antennas and receiver k has M_k antennas. Receiver k is only interested in decoding the message sent by transmitter k and treats the interference received from all other transmitters $j \neq k$ as additional noise. The channel matrices $\boldsymbol{H}_{kj} \in \mathbb{C}^{M_k \times N_j}$ contain the complex channel gains between the N_j antennas of transmitter j and the M_k antennas of receiver k. With the vector $\boldsymbol{y}_k \in \mathbb{C}^{M_k}$ containing the symbols received by user k, the noise vector $\boldsymbol{n}_k \in \mathbb{C}^{M_k}$ experienced by receiver k, and the vector of symbols $\boldsymbol{x}_j \in \mathbb{C}^{N_j}$ transmitted by user j, our system is defined by

$$\boldsymbol{y}_k = \boldsymbol{H}_{kk}\boldsymbol{x}_k + \sum_{j \neq k} \boldsymbol{H}_{kj}\boldsymbol{x}_j + \boldsymbol{n}_k \qquad \forall k \in \{1, \ldots, K\},$$

where the first summand is the desired part of the signal and the remaining terms are undesired, i. e., the interference and noise. We assume that the noise at each receiver is uncorrelated and has variance σ^2 at each antenna, i. e.,

$$\mathrm{E}\left[\boldsymbol{n}_k\boldsymbol{n}_k^{\mathrm{H}}\right] = \sigma^2\mathbf{I} \qquad \forall k \in \{1, \ldots, K\}.$$

The transmit covariance matrices are defined as

$$\boldsymbol{Q}_k = \mathrm{E}\left[\boldsymbol{x}_k\boldsymbol{x}_k^{\mathrm{H}}\right] \qquad \forall k \in \{1, \ldots, K\},$$

and are by definition positive semi-definite. The transmit signals are subject to the unit transmit power constraints

$$\mathrm{E}\big[\|\boldsymbol{x}_k\|_2^2\big] = \mathrm{tr}(\boldsymbol{Q}_k) \leq 1 \qquad \forall k \in \{1, \ldots, K\}. \tag{1}$$

In this work we consider only *single-stream transmission* or *beamforming*, meaning that the vector of symbols transmitted by user k has the form

$$\boldsymbol{x}_k = \boldsymbol{v}_k s_k$$

where $\boldsymbol{v}_k \in \mathbb{C}^{N_k}$ is the (constant) beamforming vector and $s_k \in \mathbb{C}$ is the scalar unit-variance data symbol. Therefore, the covariance matrix $\boldsymbol{Q}_k = \boldsymbol{v}_k \boldsymbol{v}_k^{\mathrm{H}}$ has at most rank one and the power contraint (1) becomes

$$\|\boldsymbol{v}_k\|_2^2 \leq 1 \qquad \forall k \in \{1, \ldots, K\}.$$

On the receiver side, the receive filter vector $\boldsymbol{g}_k \in \mathbb{C}^{M_k}$ is applied to obtain the estimate of the data symbol

$$\hat{s}_k = \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{y}_k.$$

Since the interference is treated as noise, our figure of merit for user k is the SINR:

$$\gamma_k = \frac{|\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{H}_{kk} \boldsymbol{v}_k|^2}{\sum_{j \neq k} |\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{H}_{kj} \boldsymbol{v}_j|^2 + \|\boldsymbol{g}_k\|_2^2 \sigma^2}. \tag{2}$$

Note that γ_k is invariant to multiplication of $\boldsymbol{g}_k$ by a non-zero scalar. Therefore, it is often assumed for notational convenience that $\|\boldsymbol{g}_k\|_2^2 = 1$. Also, note that for $\boldsymbol{g}_k = \mathbf{0}$ the SINR is not defined.

III. Maximization of the Sum Rate

We define an overall optimization objective by means of utility functions: each user's utility $u_k(\gamma_k)$ is an increasing function of its SINR γ_k. The goal is to maximize the overall system efficiency, which we define as the sum of all users' utilities $\sum_k u_k(\gamma_k)$. We will mostly consider the *achievable rate* utility:

$$u_k(\gamma_k) = R_k = \log(1 + \gamma_k).$$

in nats/channel use. The problem of maximizing the overall system efficiency in that case, i. e., the sum rate maximization problem, is

$$\max_{\substack{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_K \\ \boldsymbol{g}_1, \ldots, \boldsymbol{g}_K}} \sum_{k=1}^{K} R_k \quad \text{s. t.:} \quad \|\boldsymbol{v}_k\|_2^2 \leq 1 \quad \forall k \in \{1, \ldots, K\}. \tag{3}$$

This problem is non-convex and can have multiple distinct local optima. Prior work has therefore largely focused on the regime of asymptotically high signal-to-noise ratio (SNR), where $\sigma^2 \to 0$ [10], [17], [18]. In the high-SNR regime to maximize the sum rate we must determine the maximum number of users that can transmit without interference, i. e., the beams and receiver filters must satisfy the *Zero-Forcing (ZF) conditions*

$$\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{H}_{kk} \boldsymbol{v}_k \neq 0 \qquad \text{and} \qquad \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{H}_{kj} \boldsymbol{v}_j = 0 \quad \forall j \neq k. \tag{4}$$

It was shown in [10], [17], [18] that if the channels contain *i. i. d.* random elements, and every transmitter has N antennas and every receiver has M antennas, then the condition (4) can be fulfilled with probability one (w. p. 1) for all users K if and only if

$$N + M - 1 \geq K. \tag{5}$$

This condition can be extended to arbitrary numbers of antennas at each transmitter/receiver, in which case a potentially large number of inequalities must be checked. When (4) is fulfilled for all K users, and the number of interferers $K-1$ exceeds $M_k - 1$, the interference is said to be *spatially aligned* at receiver k. That is, the received signals from the $K-1$ interfering transmitters occupy a lower $(M_k - 1)$-dimensional subspace of $\mathbb{C}^{M_k}$. The DoFs of the network refers to the maximum number of non-interfering data streams that can be transmitted, and corresponds to the slope of the sum-rate curve versus logarithmic SNR at high SNRs. For the single-beam per user scenario considered here, and N antennas/M antennas per transmit/receive node, the DoFs are $N + M - 1$.

Solving the sum rate maximization problem (3) numerically, especially at high SNRs, is challenging for a few reasons. First, the number of ZF solutions to (4) grows very rapidly with the system size K, M, and N. The sum rate associated with these different solutions can vary dramatically, since the ZF conditions depend only on the cross-channels, so that the beams can be grossly misaligned with the direct channels. Hence although an algorithm that finds a random aligned solution at high SNRs achieves the available DoFs, the sum rate is typically far from optimal. (This performance variation over ZF solutions is studied in [11].) Furthermore, finding *all* solutions to the ZF conditions (4) becomes computationally infeasible for all but the smallest systems.[1]

Of course, for finite SNRs, it is not optimal to zero-force the interference. In principle, problem (3) can be solved directly with a general purpose global optimization technique; however, that is again likely to be computationally infeasible unless the system is small.[2] Furthermore, that would require an extensive exchange of channel state information among the nodes.

Another challenge is that the sum-rate objective may contain sharp peaks, corresponding to aligned solutions, making it difficult to find those solutions by a conventional (e. g., gradient) method. This is illustrated in Fig. 1, which shows a contour plot of sum rate for the case $K = 3$ and $M = N = 2$

[1]A method for enumerating the aligned strategies is given in [26], but it does not scale well to large systems.

[2]In [27], [28] a global optimization method is proposed for $K = 2$ users and $M_1 = M_2 = 1$ receive antennas. The single-receive-antenna case, however, does not require interference alignment; also, the approach does not scale well to larger scenarios.

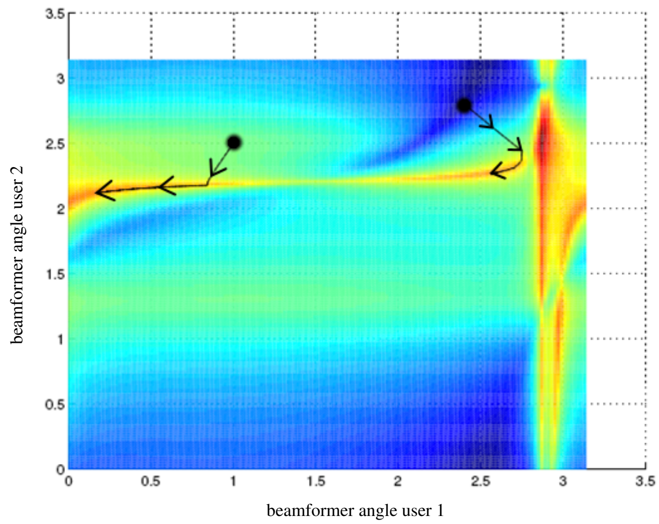


Fig. 1. Contour plot of sum rate for an example with three users, 2×2 channels, and SNR 40 dB. The contours are in a two-dimensional subspace of the beam coefficients that contains the optimal (nearly aligned) solution shown in the upper right. Also shown are trajectories of a gradient algorithm from two different initial starting points. The end point corresponds to shutting off one of the users (determined by examining the resulting beamformers).

with SNR 40 dB. The contours are shown in a two-dimensional subspace of the beam coefficients, which contains the optimal aligned solution (upper right).[3] Also shown is the trajectory of a gradient algorithm starting at two different points with three active users. In each case the algorithm finds a local optimum corresponding to two active users, i. e., it powers off one user. Hence in general other methods are needed to find aligned solutions.

Given the previous difficulties, a general goal is to find algorithms that find good (suboptimal) solutions with reasonable computational effort. We next describe a few different methods in which the beams and receivers are iteratively computed in a distributed manner, and contrast in terms of computational requirements, information exchange, and convergence.

[3]For the plot, the channels and beams were real-valued, i. e., $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3 \in \mathbb{R}^2$, so that each beam can be parameterized by one angle; the variable space is therefore a three-dimensional cube with edges of length π. The plot was generated by computing the sum rate (corresponding to a color) in each point of the cube for a fixed set of channel matrices and manually selecting a cross-section of the cube. The trajectories correspond to an unconstrained gradient ascent with a very small step size; they have been projected onto this cross-section.

IV. DISTRIBUTED ALGORITHMS

Here we describe the beamforming algorithms, which will be compared in the next section. First we distinguish between *centralized* and *distributed* algorithms. In a *centralized* computation model, the K users estimate all direct- and cross-channel gains and pass this information to a central controller by means of signaling links. The central controller then solves the optimization problem (3) directly (or finds an approximate solution) and passes the beams and receivers back to the users via the signaling links. In contrast, in a *distributed* computation model, the transmitters and receivers update associated beams and receive filters autonomously given "local" channel state information. In general, this local information must be a strict subset of all channel states.

Here we consider a set of distributed algorithms in which the receiver updates depend only on information available at the corresponding receiver and beam updates require knowledge of the direct-channel and possibly the cross-channels to neighboring receivers (including the receiver filters). Of course, that information must be exchanged prior to the corresponding updates. Starting from an initial assignment of beams and receive filters, the users iteratively update their beams and receive filters using this information.

We also distinguish between *parallel* and *sequential* update schedules. Parallel updates imply that

- all K users update their beams simultaneously;
- all K users update their receive filters simultaneously (taking into account the new beams from the preceding step);
- the users announce each update over the signaling links.

For sequential updates, the users take turns updating their beams/receive filters. Specifically,

- transmitter k updates $\boldsymbol{v}_k$;
- receiver k updates $\boldsymbol{g}_k$;
- these updates are announced to the other users.

During this update all other beams and receivers remain constant, and the users may update either according to a fixed schedule (e. g., round-robin) or asynchronously.[4]

The initial assignment of beams and receive filters affects the performance of the algorithms to be discussed. We consider two initial assignments of beams: a random assignment and an assignment, which is optimal at low SNRs. (The receivers can be subsequently optimized.) At low SNR, i. e., when

[4] See also [29], which introduces the distinction between parallel and sequential updates in the context of signature optimization for CDMA.

the noise power σ^2 is large compared to the interference terms in the denominator of the SINR (2), the SINR (equivalently, R_k) is maximized by choosing the beam $\boldsymbol{v}_k$ to be the principal eigenvector of $\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{H}_{kk}$ and $\boldsymbol{g}_k$ is the complex conjugate of the principal eigenvector of $\boldsymbol{H}_{kk}\boldsymbol{H}_{kk}^{\mathrm{H}}$. This initialization generally performs better than a random initialization, and assumes that the transmitter knows the direct channel.

In what follows we state convergence results for each algorithm. Although for some of the algorithms considered the *objective* is known to converge, the beamformers and receive filters are not proven to converge for any algorithm. Nevertheless, a scenario in which the objective converges, but the beamformers do not converge has not been observed.

A. Selfish Updates

This refers to the strategy in which each user maximizes its own SINR without considering the interference caused to the other users. Without the restriction of a single stream per user (so that $\boldsymbol{v}_k$ and $\boldsymbol{g}_k$ can become matrices), sequential selfish updates that maximize each user's rate have been studied in [23], [30]–[33]. In general, selfish update schemes cannot achieve alignment, and therefore do not achieve the available DoFs.

Objective: Each user k maximizes its own SINR γ_k.

Updates: Given $\boldsymbol{g}_k$, the updated beam is

$$\boldsymbol{v}_k^{\mathrm{new}} = \alpha \cdot \boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*$$

where α is chosen such that $\|\boldsymbol{v}_k^{\mathrm{new}}\|_2^2 = 1$ and given $\boldsymbol{v}_k$, the updated receive filter is

$$\boldsymbol{g}_k^{\mathrm{new}\,\mathrm{T}} = \beta \cdot \boldsymbol{v}_k^{\mathrm{H}}\boldsymbol{H}_{kk}^{\mathrm{H}}\left(\sum_{j\neq k}\boldsymbol{H}_{kj}\boldsymbol{v}_j\boldsymbol{v}_j^{\mathrm{H}}\boldsymbol{H}_{kj}^{\mathrm{H}} + \sigma^2\mathbf{I}\right)^{-1} \tag{6}$$

where β is chosen such that $\|\boldsymbol{g}_k^{\mathrm{new}}\|_2^2 = 1$. These updates can be sequential or parallel.

Information Exchange: The transmitter update requires knowledge of the direct channel matrix $\boldsymbol{H}_{kk}$ and the current receive filter $\boldsymbol{g}_k$. The former only needs to be fed back once before the initial iteration, while the latter is continually updated and therefore must be fed back after each iteration. The receiver update can be accomplished via standard estimation methods (e. g., with a pilot sequence). The selfish update strategy therefore does not require any information exchange between different users.

Convergence: This algorithm is not proven to converge; in fact, oscillations can be observed in numerical experiments for particular channel realizations.

B. Min-Leakage

The min-leakage algorithm was proposed in [12] as a numerical method for determining whether or not interference alignment is possible.

Objective: Each update minimizes the sum interference power

$$I_{\mathrm{sum}} = \sum_{k} \sum_{j \neq k} |\boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk} \boldsymbol{v}_k|^2.$$

Updates: $\boldsymbol{v}_k^{\mathrm{new}}$ is the eigenvector corresponding to the smallest eigenvalue of the matrix $\sum_{j \neq k} \boldsymbol{H}_{jk}^{\mathrm{H}} \boldsymbol{g}_j^* \boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk}$ and $\boldsymbol{g}_k^{\mathrm{new}}$ is the complex conjugate of the eigenvector corresponding to the smallest eigenvalue of $\sum_{j \neq k} \boldsymbol{H}_{kj} \boldsymbol{v}_j \boldsymbol{v}_j^{\mathrm{H}} \boldsymbol{H}_{kj}^{\mathrm{H}}$. The updates can be done sequentially or in parallel.

Information Exchange: To compute the preceding updates directly transmitter k must know the combined channels and receive filters $\boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk}$ for $j \neq k$. The channel matrices can be estimated and exchanged among the users initially; however, the receive filters must then be announced to the other users after each iteration. Alternatively, the beams can be computed by transmitting pilots synchronously from the receivers to the transmitters using the receive filters as beams [34]. In that way the covariance matrix $\sum_{j \neq k} \boldsymbol{H}_{jk}^{\mathrm{H}} \boldsymbol{g}_j^* \boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk}$ can be directly estimated. Similarly, the receiver update requires knowledge of the covariance matrix of the received interference, which can be estimated locally by means of pilot sequences.

Convergence: The sum interference power I_{sum} cannot be increased by an update, both with parallel and sequential schedules; therefore the objective converges.

C. Max-SINR

The max-SINR algorithm was also proposed in [12], and is motivated by uplink-duality for the multiple-access and broadcast channels. Although this duality does not apply for interference networks, the max-SINR often gives near-optimal performance.

Objective: For receiver updates, the objective is to maximize the SINR γ_k; for transmitter updates, the objective is to maximize the "reverse SINR"

$$\xi_k = \frac{|\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{H}_{kk} \boldsymbol{v}_k|^2}{\sum_{j \neq k} |\boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk} \boldsymbol{v}_k|^2 + \sigma^2}.$$

This corresponds to reversing the direction of transmission so that the roles of the beams and receive filters are swapped. The objective ξ_k is then the SINR in the reverse direction (at transmitter k). Note that the denominator of the reverse SINR ξ_k contains the interference *caused* by transmitter k instead of the interference experienced by receiver k.[5]

[5] This objective is also considered in [35] for $M_k = 1$ antenna at each receiver.

Updates: The transmitter/receiver updates can be done sequentially or in parallel. (The simulation results assume parallel updates.) Given the set of receive filters $\boldsymbol{g}_1, \cdots, \boldsymbol{g}_K$, the beam updates are

$$\boldsymbol{v}_k^{\text{new}} = \alpha \cdot \left(\sum_{j \neq k} \boldsymbol{H}_{jk}^{\text{H}} \boldsymbol{g}_j^* \boldsymbol{g}_j^{\text{T}} \boldsymbol{H}_{jk} + \sigma^2 \mathbf{I} \right)^{-1} \boldsymbol{H}_{kk}^{\text{H}} \boldsymbol{g}_k^* \tag{7}$$

where α is chosen such that $\|\boldsymbol{v}_k^{\text{new}}\|_2^2 = 1$. Given the set of beams $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_K$, the receiver updates are given by (6).

Information Exchange: To compute the update (7) directly, transmitter k must have knowledge of the combined cross-channels and receive filters $\boldsymbol{g}_j^{\text{T}} \boldsymbol{H}_{jk}$, $j \neq k$. Alternatively, as for the interference leakage algorithm, the beams can be directly estimated by transmitting pilots synchronously from the receivers in the reverse direction to the transmitters [34]. The receiver update can be accomplished by standard estimation methods.

Convergence: Whether or not the preceding parallel updates converge to a fixed-point from an arbitrary starting point is unknown, although numerical experiments indicate that convergence is quite reliable for randomly chosen channel matrices with *i. i. d.* elements.[6]

D. Minimum Mean Squared Error (MMSE)

The MMSE criterion has been proposed for optimizing transmit beams in [16] for cellular networks and in [37] for an interference network. (Analogous results for optimizing signatures in a Code-Division Multiple Access (CDMA) network were presented in [29]. See also [38], [39].) The relation between MMSE updates and max-SINR updates was presented in the conference version of this paper [15], where it was observed that MMSE updates can achieve interference alignment in interference networks. Here we give a more complete performance comparison of MMSE and related algorithms.

Objective: Each update minimizes the sum mean squared error (MSE). Letting $\varepsilon_k = \mathrm{E}[|\hat{s}_k - s_k|^2]$, this objective is

$$\sum_k \varepsilon_k = \sum_k \left(\sum_j |\boldsymbol{g}_k^{\text{T}} \boldsymbol{H}_{kj} \boldsymbol{v}_j|^2 - 2\,\mathrm{Re}\{\boldsymbol{g}_k^{\text{T}} \boldsymbol{H}_{kk} \boldsymbol{v}_k\} + \|\boldsymbol{g}_k\|_2^2 \sigma^2 \right) + K$$

and the minimization is subject to the beamformer power constraints $\|\boldsymbol{v}_k\|_2^2 \leq 1$ for all k, and no constraints on the receive filters.

[6] It is shown in [36] that if the algorithm is initialized sufficiently close to a local optimum at sufficiently high SNR, then it converges exponentially.

Updates: The updates can be parallel or sequential. For the beam update at transmitter k, first compute

$$\boldsymbol{v}_k^{\mathrm{tmp}} = \left(\sum_j \boldsymbol{H}_{jk}^{\mathrm{H}} \boldsymbol{g}_j^* \boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk}\right)^+ \boldsymbol{H}_{kk}^{\mathrm{H}} \boldsymbol{g}_k^*.$$

If $\|\boldsymbol{v}_k^{\mathrm{tmp}}\|_2^2 \leq 1$, then $\boldsymbol{v}_k^{\mathrm{new}} = \boldsymbol{v}_k^{\mathrm{tmp}}$; otherwise,

$$\boldsymbol{v}_k^{\mathrm{new}} = \left(\sum_j \boldsymbol{H}_{jk}^{\mathrm{H}} \boldsymbol{g}_j^* \boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk} + \lambda \mathbf{I}\right)^{-1} \boldsymbol{H}_{kk}^{\mathrm{H}} \boldsymbol{g}_k^*$$

where $\lambda > 0$ is chosen such that $\|\boldsymbol{v}_k^{\mathrm{new}}\|_2^2 = 1$. The regularization factor λ can be found efficiently with a line search using Newton iterations: starting with $\lambda = 0$, the regularization factor is updated according to

$$\lambda^{\mathrm{new}} = \lambda^{\mathrm{old}} + \frac{\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{H}_{kk} \left(\sum_j \boldsymbol{H}_{jk}^{\mathrm{H}} \boldsymbol{g}_j^* \boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk} + \lambda^{\mathrm{old}} \mathbf{I}\right)^{-2} \boldsymbol{H}_{kk}^{\mathrm{H}} \boldsymbol{g}_k^* - 1}{2 \boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{H}_{kk} \left(\sum_j \boldsymbol{H}_{jk}^{\mathrm{H}} \boldsymbol{g}_j^* \boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk} + \lambda^{\mathrm{old}} \mathbf{I}\right)^{-3} \boldsymbol{H}_{kk}^{\mathrm{H}} \boldsymbol{g}_k^*}$$

until convergence is achieved, which usually only requires a few iterations. The receiver update is given by

$$\boldsymbol{g}_k^{\mathrm{new T}} = \boldsymbol{v}_k^{\mathrm{H}} \boldsymbol{H}_{kk}^{\mathrm{H}} \left(\sum_j \boldsymbol{H}_{kj} \boldsymbol{v}_j \boldsymbol{v}_j^{\mathrm{H}} \boldsymbol{H}_{kj}^{\mathrm{H}} + \sigma^2 \mathbf{I}\right)^{-1}. \tag{8}$$

These updates are similar to those for the max-SINR algorithm. The receiver update is the same except that the filter is not scaled to have unit norm. (Although the sum inside the matrix inverse in (8) is taken over all j, as opposed to $j \neq k$ for the max-SINR algorithm, application of the matrix-inversion lemma shows that this amounts to a scale factor.) For the transmitter update, the inverse is regularized with the Lagrange multiplier λ instead of σ^2 in the max-SINR algorithm. The simulation results in the next section indicate that these differences can lead to significant differences in sum rates at high SNRs.

Information Exchange: As for the max-SINR algorithm, the transmitter update for user k requires knowledge of the combined channel matrices/receive filters $\boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk}$ for all j. The receiver updates can be performed with local information.

Convergence: The sum MSE is nonincreasing with each update (both for parallel and sequential update schedules) and therefore converges to a local minimum.

E. Adaptively Weighted MMSE

The weighted sum MSE criterion for beam updates has been introduced in [7], [15], [16]. In [16] an equivalence between the weighted sum rate and weighted sum MSE criteria is shown for the broadcast channel. That was extended to interfering MIMO broadcast channels in [7].[7] Here we give conditions on the user utility function that guarantee convergence with a single beam per user for the interference channel.

Objective: Each update minimizes the weighted MSE

$$\sum_k w_k \varepsilon_k$$

where the weights w_k can either be fixed or updated during the course of the iterations. This objective has the attractive property that by varying the weights, it can achieve different points in the rate region, corresponding to different user priorities. This is not possible with the min-leakage or max-SINR algorithms. Furthermore, by adapting the weights, the objective can be approximately matched to different utility objectives. (See also [16], which considers only the weighted sum rate objective).[8]

Updates: The transmitter update follows from the unweighted MMSE algorithm where each channel matrix $\boldsymbol{H}_{jk}$ is replaced by $\sqrt{w_j/w_k}\boldsymbol{H}_{jk}$, i.e., first

$$\boldsymbol{v}_k^{\mathrm{tmp}} = \left(\sum_j \frac{w_j}{w_k} \boldsymbol{H}_{jk}^{\mathrm{H}} \boldsymbol{g}_j^* \boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk}\right)^+ \boldsymbol{H}_{kk}^{\mathrm{H}} \boldsymbol{g}_k^*$$

is computed, and if $\|\boldsymbol{v}_k^{\mathrm{tmp}}\|_2^2 \leq 1$, then $\boldsymbol{v}_k^{\mathrm{new}} = \boldsymbol{v}_k^{\mathrm{tmp}}$; otherwise,

$$\boldsymbol{v}_k^{\mathrm{new}} = \left(\sum_j \frac{w_j}{w_k} \boldsymbol{H}_{jk}^{\mathrm{H}} \boldsymbol{g}_j^* \boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk} + \lambda \mathbf{I}\right)^{-1} \boldsymbol{H}_{kk}^{\mathrm{H}} \boldsymbol{g}_k^* \tag{9}$$

where $\lambda > 0$ is chosen such that $\|\boldsymbol{v}_k^{\mathrm{new}}\|_2^2 = 1$. The receiver update is the same as in the unweighted MMSE algorithm.

We can now adapt the weights to maximize the sum utility objective $\sum_{k=1}^K u_k(\gamma_k)$. With MMSE receive filters given by (8), we note that the SINR $\gamma_k = 1/\varepsilon_k - 1$, so that we can express the utility function in terms of the MSE, i.e.,

$$\bar{u}_k(\varepsilon_k) = u_k\left(\frac{1}{\varepsilon_k} - 1\right). \tag{10}$$

[7]The algorithm in [7] extends the algorithm presented here to the scenario in which there are multiple beams per user. Also, a different update schedule is used in [7]. There the objective weights are updated after each iteration whereas here those weights are updated after convergence of the inner loop of beamformer updates. We have found that the latter schedule generally gives better performance.

[8]The max-SINR updates can be modified to mimic the weighted-MSE updates thereby achieving different points near (possibly on) the boundary of the rate region. However, for the adaptively weighted MSE algorithm the inclination of a plane that "touches" the rate region at a point obtained by the algorithm is simply determined by the weights.

Expanding the sum utility in a Taylor expansion around the operating point $\varepsilon_{k,0}$ and dropping all but the linear term gives

$$\sum_{k=1}^{K} \bar{u}_k(\varepsilon_k) = \sum_{k=1}^{K} -\alpha_k \varepsilon_k + C + O(\varepsilon_k^2) \tag{11}$$

where C does not depend on any ε_k and

$$\alpha_k = -\left.\frac{\partial \bar{u}_k(\varepsilon_k)}{\partial \varepsilon_k}\right|_{\varepsilon_k = \varepsilon_{k,0}}. \tag{12}$$

For the rate utility $u_k(\gamma_k) = \log(1+\gamma_k)$ we have $\bar{u}_k(\varepsilon_k) = -\log(\varepsilon_k)$, so that the sum rate behaves locally as weighted sum MSE with weights $w_k = \alpha_k = 1/\varepsilon_{k,0}$.

The resulting *Adaptively Weighted-MSE (AW-MSE)* algorithm follows:

1) Initialize the beamformers $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_K$ arbitrarily, and compute the optimal receive filters $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_K$ from (8) and the weights $w_1, \ldots, w_K$ to be $\alpha_1, \ldots, \alpha_K$ from (12).
2) Iteratively update the beamformers from (9) and the receivers from (8) until convergence.
3) Update the weights $w_1, \ldots, w_K$ according to the new operating point $\varepsilon_{1,0}, \ldots, \varepsilon_{K,0}$ using (12).
4) Repeat from 2) until the weights $w_1, \ldots, w_K$ have converged.

When the algorithm has converged, clearly (6) (with $\beta = 1$) and (9) are fulfilled for all k. Furthermore, since the weights w_k correspond to the current value of ε_k, the necessary optimality conditions for maximizing sum utility are fulfilled. Hence if the algorithm converges, it finds a locally optimal solution. A similar algorithm has been proposed for MIMO broadcast channels in [16].

Information Exchange: The AW-MSE algorithm requires the same information as the unweighted MMSE and max-SINR algorithms. Additionally, the weights w_k must be exchanged among the users whenever they are updated.

Convergence: Convergence of the sum utility objective requires the following constraints on the utility functions, where u_k' and u_k'' denote the first- and second-derivatives of $u_k(\cdot)$.

Proposition 1. *If u_k satisfies $(\gamma_k + 1)u_k'' + 2u_k' \geq 0$, or equivalently, $\bar{u}_k'' \geq 0$ for all feasible γ_k or ε_k and for each k, then the sum utility given by the AW-MSE algorithm converges.*

Proof: The proof consists of showing that updating the beamformers and receive filters, given new (updated) weights, will increase the sum utility relative to that before the last weight update. This is true because the condition $\bar{u}_k'' \geq 0$ implies that the utility function is a convex function of the MMSE, so that the linearization of the objective in (11) lower bounds the sum utility objective and is tight at the current operating point, i. e.,

$$\sum_{k=1}^{K} \bar{u}_k(\varepsilon_k) \geq \sum_{k=1}^{K} \bar{u}_k(\varepsilon_{k,0}) - \sum_{k=1}^{K} w_k(\varepsilon_k - \varepsilon_{k,0}) \tag{13}$$

where $\varepsilon_{k,0}$ is user k's MSE at the current operating point.

Given a set of new weights, w_k's, the users then update their beamformers and receive filters according to step 2. After those updates we must have

$$\sum_{k=1}^{K} w_k \varepsilon_{k,*} \leq \sum_{k=1}^{K} w_k \varepsilon_{k,0}. \tag{14}$$

where $\varepsilon_{k,*}$ is the updated MSE for user k. Combining (14) and (13) implies that

$$\sum_{k=1}^{K} \bar{u}_k(\varepsilon_{k,*}) \geq \sum_{k=1}^{K} \bar{u}_k(\varepsilon_{k,0}) \tag{15}$$

or equivalently,

$$\sum_{k=1}^{K} u_k(\gamma_{k,*}) \geq \sum_{k=1}^{K} u_k(\gamma_{k,0}), \tag{16}$$

which means the sum utility cannot decrease after the beamformers and receive filters are updated. Since the sum utility is bounded, it must therefore converge. ■

The proof applies when the weights, beamformers, and receive filters are updated asynchronously, as long as the receive filters are optimized for the current set of beams before each weight update. That is, (14) still holds, which guarantees that the sum utility cannot decrease over consecutive weight updates. The condition in Proposition 1 applies to the rate utility $\log(1+\gamma)$, but excludes, e. g., the α-fair utility γ^α/α. Although the preceding proof implies convergence of the sum utility objective, and not the beamformers, convergence of the beams is always observed in simulations.

F. Interference Pricing

Objective: Interference pricing was introduced in [14] to maximize the sum utility objective $\sum_{k=1}^{K} u_k(\gamma_k)$ over transmit powers in single-antenna interference networks. Extension to multi-antenna networks is discussed in [8] and the references therein.

Updates: For MIMO interference networks interference pricing consists of the following three types of updates, which can be performed asynchronously:

1) Each receiver j announces an *interference price* to all transmitters,

$$\pi_j = -\frac{\partial u_j(\gamma_j)}{\partial I_j}, \tag{17}$$

which is the marginal decrease in utility for an increase in received interference $I_j = \sum_{i \neq j} |\boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{ji} \boldsymbol{v}_i|^2$.

2) Each transmitter updates its beam to maximize a *best response* objective, which is its own utility minus the "cost" of the interference it produces, i. e., it solves

$$\max_{\boldsymbol{v}_k} \; u_k(\gamma_k) - \sum_{j \neq k} \pi_j |\boldsymbol{g}_j^{\mathrm{H}} \boldsymbol{H}_{jk} \boldsymbol{v}_k|_2^2 \qquad \text{s. t:} \quad \|\boldsymbol{v}_k\|_2^2 \leq 1. \tag{18}$$

The right (cost) term is a linearization of the sum utility excluding user k.

3) Each receiver updates its filter according to (6), which can be locally estimated. We will assume that β is chosen such that $\|\boldsymbol{g}_k^{\mathrm{new}}\|_2^2 = 1$.

For the sum rate utility the interference price is

$$\pi_j = \frac{1}{\sum_{i \neq j} |\boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{ji} \boldsymbol{v}_i|^2 + \sigma^2} - \frac{1}{\sum_i |\boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{ji} \boldsymbol{v}_i|^2 + \sigma^2}. \tag{19}$$

The solution to (18) is in general not available in closed form, but can be found by means of a bisection line search. A description of the procedure is given in the Appendix; the derivation can be found in [40].[9] A stationary point of this algorithm can be shown to fulfill the Karush-Kuhn-Tucker (KKT) conditions for maximizing the sum utility objective.

Information Exchange: As for the max-SINR and MMSE algorithms, the transmitter update for user k requires knowledge of the combined channel matrices/receive filters $\boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk}$ for all j. The best-response update also depends on the interference prices π_j for all j, but only through the product $\sqrt{\pi_j} \boldsymbol{g}_j^{\mathrm{T}} \boldsymbol{H}_{jk}$, which can be exchanged directly.

In addition to updating the beam, the transmitter must update the interference price π_k according to (19). It is straightforward to show that π_k can be computed from the current SINR γ_k as well as $|\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{H}_{kk} \boldsymbol{v}_k|^2$; therefore both $\boldsymbol{g}_k^{\mathrm{T}} \boldsymbol{H}_{kk}$ and γ_k must be fed back from receiver k to transmitter k after each update. While $\boldsymbol{g}_k$ is needed for the transmitter update in all of the algorithms considered, γ_k is not explicitly needed for those updates (although σ^2 may be required). Nevertheless, to choose a coding scheme for the payload data, the transmitter must know the channel quality at its associated receiver, so that we implicitly assume that γ_k is fed back from receiver k to transmitter k for all algorithms.

Convergence: With sequential updates and up-to-date prices, i. e., all prices are updated and exchanged every time a beamformer changes, the sum rate is non-decreasing and therefore converges [42]. Additional convergence results for more general utility functions are summarized in [8]. Numerical experiments suggest that convergence with different (e. g., parallel) update schedules is also reliable. Furthermore,

[9] A simplified version of the pricing algorithm, in which an approximation of (18) is solved for the transmitter update, was presented in [41] for MISO channels. The simplified version does not require a line search for the transmitter update, but potentially does not converge.

convergence to a fixed point implies that the KKT conditions of the sum utility maximization problem are fulfilled.

G. Pricing with Incremental SNR

Objective: We now present another method for maximizing sum utility, which attempts to track a local optimum as the SNR is incrementally increased from zero. (Equivalently, all transmitters increase their power simultaneously.) This is motivated by the observation that as the SNR tends to zero, the noise dominates the interference, so that the optimal beam $\boldsymbol{v}_k$ tends to the principal eigenvector of the direct channel $\boldsymbol{H}_{kk}^H\boldsymbol{H}_{kk}$. Furthermore, as the SNR tends to infinity, the set of local optima correspond to ZF (interference aligned) solutions (which depend only on the cross-channel matrices). The set of beams that maximize the sum utility at high SNRs should therefore be close to a ZF solution where the beams are as closely aligned as possible to the low-SNR solution (which depends only on the direct channels).

This intuition is illustrated in Fig. 2, where, for illustrative purposes, it is assumed that the optimization problem is one-dimensional. (The "beam coefficients" represent $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_K$ assuming that the corresponding optimal receivers are used). When the SNR is low, the sum utility is uniquely maximized by beams that are (approximately) aligned to the principal modes of the direct channels. As the SNR increases to the point where interference dominates, many additional local optima appear including ZF solutions for beamformers/receivers as well as possibly other local optima, corresponding to solutions where the interference is zero for some subset of users. (This is supported by the observation that gradient algorithms that follow the steepest ascent of the sum rate occasionally fail to find an interference aligned solution at high SNRs.) The locations of the local optima are determined by the cross-channels.

Updates: We fix the transmit power constraints and define a decreasing sequence of noise powers that ends with the actual noise power σ^2:

$$\sigma_{(1)}^2 > \sigma_{(2)}^2 > \ldots > \sigma^2.$$

(Alternatively, we could fix the noise variance and increase the transmit power constraints.) The beams are initialized as the low-SNR solution. Starting with the largest noise power $\sigma_{(1)}^2$, the beamformers/receive filters are then updated according to the interference pricing algorithm.[10] Once the beams/receive filters have converged or a maximum number of iterations is reached, the updates are continued with noise vari-

[10]The incremental SNR approach could be combined with other algorithms as well. Here interference pricing is used, since it is a gradient-based approach, and the gradients (prices) incrementally adjust as the SNR increases.

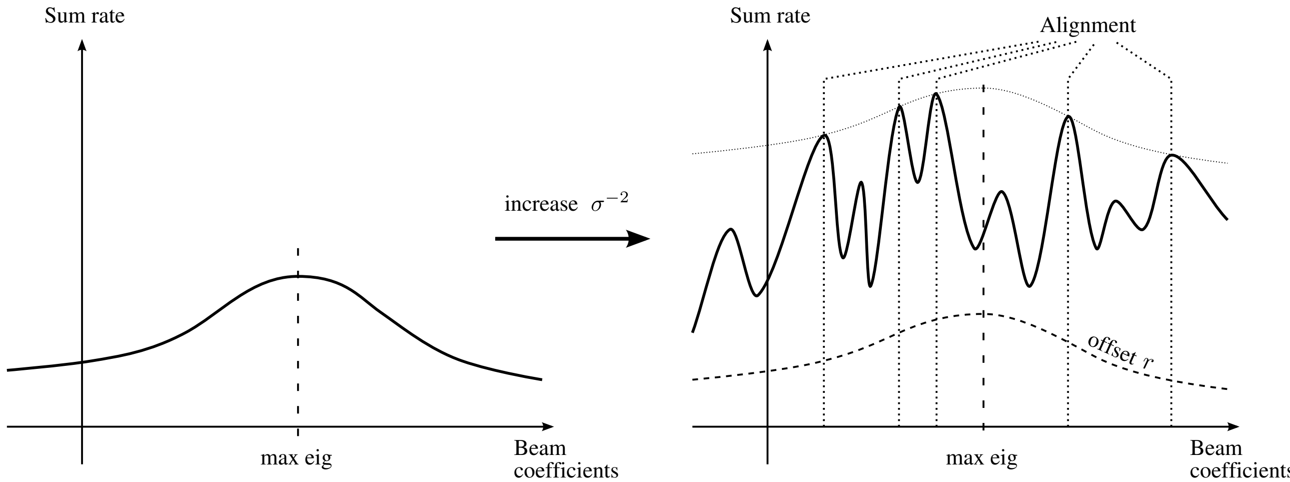


Fig. 2. For low SNR (left) the optimal beamformers are the principal eigenvectors of the Gramians of the direct channels. At high SNR (right), there are many local optimizers corresponding to ZF (aligned) solutions, which depend only on the cross channels.

ance $\sigma^2_{(2)}$. Upon convergence or a maximum number of iterations, the noise variance is again incremented and this procedure continues until the noise variance is at the final value.

Information Exchange: The information exchange requirements are identical to the interference pricing algorithm; some additional signaling may be necessary to determine when to increment the SNR (transmitted power).

Convergence: Since the updates are identical to the interference pricing algorithm, the convergence properties are also the same.[11]

H. Summary of Properties

Table I summarizes some of the properties discussed for the preceding algorithms. The columns indicate whether the power can be strictly less than the constraint, whether the algorithm is designed to find a local optimum for maximizing sum-rate obective in (3), whether convergence is guaranteed in some sense, and what sort of computations are necessary for each user's update. As previously indicated, the feedback and signaling requirements for all algorithms considered (except selfish updates) are very similar.

A count of the number of flops per update, based on results in [43], shows that all receiver updates are $O(M^3 + KM^2 + KMN)$ and all transmitter updates are $O(N^3 + KN^2 + KMN)$ with the exception of

[11]The incremental SNR algorithm was proposed in [9], but with the simplified version of the pricing updates, which is not guaranteed to converge.

TABLE I

ALGORITHM PROPERTIES

Algorithm	Power control	Sum rate objective	Convergence proven	Computation per iteration
Selfish updates	No	No	No	Solve linear eq. system
Min-leakage	No	No	Yes	Eigen-Decomposition (EVD)
Max-SINR	No	No	Locally	Solve linear eq. system
MMSE	Yes	No	Yes	Linear eq. system, line search
Weighted MMSE	Yes	Yes	Yes	Linear eq. system, line search
Pricing	Yes	Yes	For seq. updates	EVD, line search
Incremental SNR	Yes	Yes	For seq. updates	EVD, line search

the selfish transmitter update, which is $O(MN)$. For the transmitter updates that contain an inner loop, i. e., the pricing and MMSE based algorithms, each iteration of the inner loop is $O(N^3)$. Although this suggests that the algorithms have similar complexity, the associated constant factors differ substantially. For example, the constant for computing an eigenvector is far greater than that for solving a set of linear equations, and depends on precision requirements in addition to algorithmic specifications. Furthermore, the number of updates needed to achieve a target performance may also vary with the system size.

I. Other Algorithms

Here we briefly mention other algorithms, which have been proposed for beamformer optimization, and which are not included in our comparison. In [19], e. g., a min-leakage update step is combined with a step in the direction of the sum-rate gradient. Because this finds ZF solutions, it does not generally perform well at moderate and low SNRs; also, the performance strongly depends on how the gradient step-size is adjusted over time. Similarly, the performance of the gradient ascent method proposed in [20] is sensitive to the selection of a step-size and can exhibit convergence problems. In [21], the min-leakage objective is modified with an additional summand that rewards a high desired signal power; the performance again depends on a weighting factor that must be found by trial and error, and appears to be strongly dependent on the noise power. In [22], an algorithm is proposed that maximizes the “global SINR” defined as the ratio of the sum of all users’ desired powers to the sum of all users’ received interference plus noise powers; we do not include this in the comparison since its performance and convergence behavior was observed to be almost identical to that of the max-SINR algorithm, and it requires additional information exchange.

In [44], an MMSE-based algorithm is proposed with a *sum* power constraint at the transmitters as opposed to the individual power constraints assumed here. Finally, in [23] and [24] algorithms are proposed that sequentially update the transmit covariance matrices $\boldsymbol{Q}_k$ (selecting the rank as well), whereas here we constrain those matrices to be rank-one.

V. Performance Comparisons

We now present numerical comparisons of sum-rate and convergence behavior for the preceding algorithms in an interference network.

A. Simulation Parameters

The elements of the direct- and cross-channels are independent circularly symmetric complex Gaussian random variables with mean zero. The variance of the direct-channel elements is one. For the cross-channels we examine two cases: “strong cross-channels” have elements with variance one, and “weak cross-channels” have elements with variance 0.01.

We simulate the algorithms with parallel updates, as discussed in the beginning of Section IV. Further numerical experiments indicate that the performance is quite similar to that with sequential updates. We initialize all algorithms with the low-SNR optimal solution, which we refer to as “Maximum Eigenvector (ME)”; further experiments show that this provides a significant advantage over a random initialization for all algorithms.

All algorithms are run until convergence or a maximum number of iterations i_{max} is reached. (One iteration is counted when all users have updated their beamformers and receive filters.) The convergence criterion is $\sum_k \|\boldsymbol{v}_k^{\mathrm{new}} - \boldsymbol{v}_k^{\mathrm{old}}\|_2 < \epsilon$, where the threshold ϵ is set experimentally for each algorithm. Specifically, for the min-leakage, max-SINR, and incremental SNR algorithms, $\epsilon = 10^{-4}$ and $i_{max} = 10\,000$; a smaller ϵ does not lead to a visible performance improvement, whereas a larger ϵ does appear to have a negative effect. The MMSE-based algorithms and the pricing algorithm (without incremental SNR), on the other hand, exhibit a significant performance degradation at high SNRs with these parameters; we therefore set $\epsilon = 10^{-5}$ and $i_{max} = 100\,000$ for these three algorithms. Here our objective is to illustrate performance assuming the algorithms can run for a very long time.[12] Subsequently we will present results fixing i_{max} for all algorithms, cf. Fig. 5.

[12] As an exception, the results in Fig. 4a (and Table III (left)) indicate that the AW-MSE algorithm might benefit from allowing more iterations at high SNRs. We chose to keep $i_{max} = 100\,000$ to limit the computational complexity. The same holds for the pricing algorithm, which in some scenarios appears to also require more than 100 000 iterations to converge.

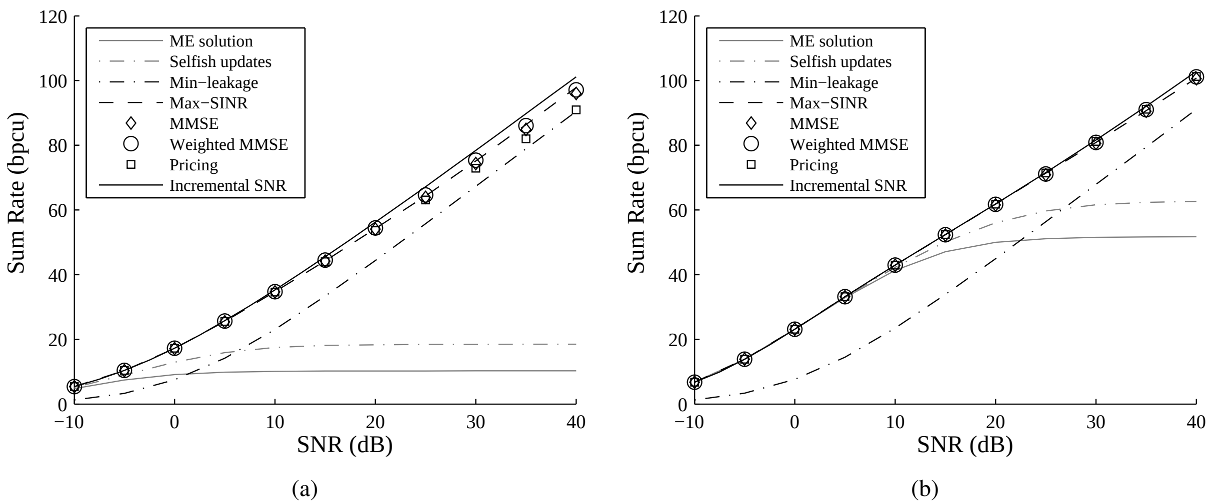


Fig. 3. Comparison of sum rate vs SNR for an interference network with $K = 7$ users and $N = M = 4$ antennas at each transmitter and receiver. The left and right plots correspond to strong and weak cross-channels, respectively. Results are averaged over 100 channel realizations.

For the MMSE-based algorithms the line search to enforce the power constraint was performed with 10 Newton iterations; for the pricing updates the bisection line search was performed with 20 iterations.[13] For the incremental SNR algorithm increments of 2.5 dB were used with the initial noise power $\sigma^2_{(1)} = 10$. At high SNRs the performance can be sensitive to both the SNR increments and the convergence threshold ϵ.

B. Sum Rate vs SNR Comparisons

Figs. 3a and 3b show sum rate versus SNR (inverse noise power σ^{-2}) in dB obtained from running the algorithms until convergence with $K = 7$ users and $N = M = 4$ antennas at each transmitter and receiver. The sum rate is averaged over 100 channel realizations. Figs. 3a and 3b show results for strong and weak cross-channels, respectively. Note that the system dimensions fulfill the ZF condition (5) with equality, i. e., it is possible to achieve zero interference for all users, but with an additional user interference alignment becomes infeasible. In Table II we show the median number of iterations required for convergence by the algorithms for low, moderate, and high SNR. In parentheses is the number of channel realizations for which the convergence criterion is not satisfied before the maximum number of iterations is reached. The iteration number for the incremental SNR algorithm includes iterations for all increments up to the given SNR value.

[13] Bisection is used instead of Newton iterations since the slope information is assumed to be unavailable.

TABLE II

MEDIAN NUMBER OF ITERATIONS FOR FIG. 3A (LEFT) AND FIG. 3B (RIGHT).

Algorithm	–10 dB	15 dB	40 dB	–10 dB	15 dB	40 dB
Selfish updates	22 (0)	51 (10)	51.5 (9)	11 (0)	30 (0)	51.5 (14)
Min-leakage	1066.5 (0)	1066.5 (0)	1066.5 (0)	1270.5 (0)	1270.5 (0)	1270.5 (0)
Max-SINR	36.5 (0)	792.5 (0)	1009 (0)	12 (0)	69 (0)	1150.5 (0)
MMSE	52 (0)	1258.5 (0)	27 441 (1)	27 (0)	383.5 (0)	25 010 (0)
Weighted MMSE	136.5 (0)	4558 (0)	29 690.5 (3)	28 (0)	1064.5 (0)	38 311.5 (4)
Pricing	51 (0)	1698 (0)	100 000 (66)	18 (0)	96.5 (0)	2982 (0)
Incremental SNR	39.5 (0)	2140.5 (0)	6242 (12)	12 (0)	292.5 (0)	3363.5 (0)

In parentheses is the percentage of channel realizations for which the algorithm did not converge before reaching the maximum iteration number.

We observe that selfish updates provide an improvement over the ME initial condition, but that the performance saturates at high SNR, since users are not penalized for causing interference. Also, selfish updates fail to converge in some cases, as pointed out in [33]. The other algorithms achieve roughly the same slope at high SNR, which indicates that they achieve interference alignment. Because the min-leakage algorithm does not take into account the direct-channels or the noise power, it suffers an SNR loss relative to the max-SINR, MMSE, and pricing algorithms. The performance of the interference pricing algorithm deteriorates slightly at high SNR for the scenario with strong cross-channels. This can be explained by examining Table II, which indicates that the pricing algorithm has difficulty converging to a stationary point even within 100 000 iterations. The incremental SNR heuristic alleviates the convergence problems and leads to the best performance of all algorithms. The max-SINR and MMSE-based algorithms achieve similar sum-rate performance, but the MMSE algorithm requires many more iterations for convergence at high SNR.

Figs. 4a and 4b show results for an interference network with $K = 10$ users and $N = M = 4$ antennas; the corresponding median iteration numbers are shown in Table III. Here the feasibility condition for interference alignment (5) is not fulfilled, i. e., optimal sum-rate performance at high SNR requires individual users to be deactivated. Clearly, this puts the algorithms that do not have the capability to control the power of the beamformers (cf. Table I) at a disadvantage: the corresponding sum rates saturate at high SNR. While the unweighted MMSE algorithm is capable of deactivating users, it does not allocate the power according to the sum rate objective, which leads to a suboptimal slope in Fig. 4a.

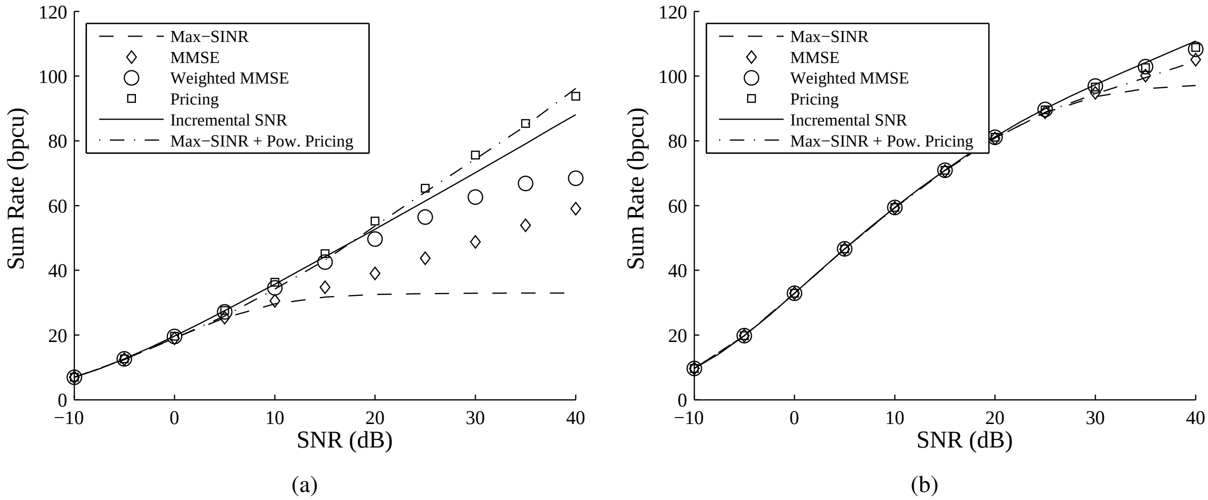


Fig. 4. Comparison of sum rate vs SNR for an interference network with $K = 10$ users and $N = M = 4$ antennas at each transmitter and receiver. The left and right plots correspond to strong and weak cross-channels, respectively. Results are averaged over 100 channel realizations.

TABLE III

MEDIAN NUMBER OF ITERATIONS FOR FIG. 4A (LEFT) AND FIG. 4B (RIGHT).

Algorithm	–10 dB	15 dB	40 dB	–10 dB	15 dB	40 dB
Max-SINR	45 (0)	243.5 (0)	255.5 (0)	13 (0)	93 (0)	225 (0)
MMSE	63 (0)	628 (0)	16 398 (0)	28 (0)	395 (0)	8942.5 (0)
Weighted MMSE	182.5 (0)	9731.5 (0)	100 000 (99)	29 (0)	1478.5 (0)	100 000 (79)
Pricing	68 (0)	1180 (0)	100 000 (52)	20 (0)	129.5 (0)	2135.5 (0)
Incremental SNR	52.5 (0)	2074 (0)	5724.5 (9)	14 (0)	390.5 (0)	2842.5 (0)
Max-SINR + Power Pricing	49 (0)	1845.5 (0)	3997 (0)	15 (0)	90.5 (0)	4312.5 (0)

In parentheses is the percentage of channel realizations for which the algorithm did not converge before reaching the maximum iteration number.

The MMSE algorithm with adaptive weights, on the other hand, has severe difficulty converging within 100 000 iterations at high SNR. The pricing and incremental SNR algorithms perform best, with the incremental SNR heuristic again leading to better convergence properties. Also shown is the sum rate achieved by the max-SINR algorithm as an inner loop with interference pricing used to adapt user powers

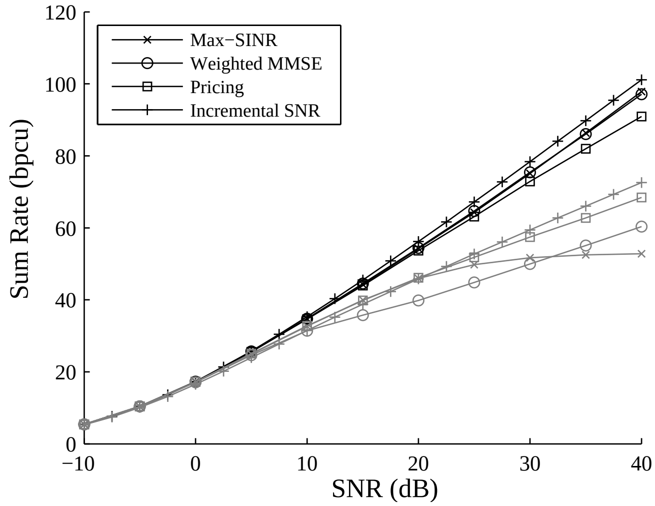


Fig. 5. Comparison of sum rates vs SNR with $K = 7$ users, $N = M = 4$ antennas, and strong cross-channels. Two sets of curves are shown corresponding to 100 iterations (shown as dark) and 20 iterations (shown as light). Results are averaged over 100 channel realizations.

as an outer loop.[14] In that way, the outer loop can deactivate users. Fig. 4a shows that on average this algorithm along with pricing performs the best among the algorithms considered.

We observe that the asymptotic slope achieved by the max-SINR plus pricing algorithm is about 20 bpcu per 10 dB, or $6.0 \log_2 10$, corresponding to six DoFs on average. Note that the maximum achievable DoFs is seven. Hence this suggests that rather than rely on adaptive power control to optimize the number of users, at high SNRs better performance can be achieved by using admission control (a scheduler) to pre-select an active set of users with fixed powers.

Fig. 5 shows results for the scenario in Fig. 3a with a limit of 100 iterations (shown as dark) and 20 iterations (shown as light) for each algorithm. For the adaptive MMSE algorithm the weights are updated after each iteration; for the incremental SNR algorithm we permit five iterations (one iteration for the lighter curves) per SNR value, so that for an SNR of 40 dB, 105 (21) iterations are performed, whereas at 30 dB only 85 (17) iterations are performed. Comparison of the sum rates in Fig. 5 with those in Fig. 3a, where the algorithms are run until convergence, shows that allowing only 100 iterations incurs a small performance loss. The curves for 20 iterations show a large performance loss relative to 100 iterations beyond about 15 dB. The performance of the max-SINR saturates at about 25 dB, whereas the

[14] The maximum number of iterations is 10^4, and the power is updated when the beams in the inner (max-SINR) loop have converged.

other algorithms exhibit positive slope (DoFs) at high SNRs.

C. Discussion

The preceding results show that the min-leakage and max-SINR algorithms generally exhibit the most reliable and fastest convergence. The sum rate achieved by the max-SINR algorithm is significantly greater than that achieved by the min-leakage algorithm at low to moderate SNRs, and is close to the maximum over the algorithms considered provided that interference alignment is feasible. However, because min-leakage and max-SINR assume fixed powers, they become interference limited when interference alignment is infeasible. In contrast, the incremental SNR algorithm in general converges more slowly and has a more complex update procedure, but can deactivate users; hence it can, in principle, achieve the maximum DoFs. It is also observed to achieve a slightly higher sum rate than the max-SINR algorithm in the feasible scenario. This may be attributed to the fact that the incremental SNR algorithm, in contrast to the max-SINR algorithm, attempts to maximize directly the sum-rate objective.

Although the MMSE algorithms offer comparable performance with the max-SINR algorithm at low to moderately high SNRs, at very high SNRs the MMSE filters converge much more slowly. This appears to be a consequence of the property that at high SNRs, the MSE objective does not sufficiently penalize small, but non-zero, interference. Consequently, the MSE-based algorithms provide very small improvements from one iteration to the next. Still, the weighted sum MSE criterion may be useful in practice in conjunction with a scheduler that determines the weights to set user priorities.[15] The convergence of the pricing algorithm without the incremental SNR heuristic also becomes problematic at high SNRs. Here, however, the difficulty is that each update results in a change in beamformers that is too large rather than too small; due to the irregularity of the sum rate objective at high SNR, the beamformers must be relatively close to a local optimum in order to be reliably attracted. The incremental SNR algorithm can alleviate this problem by ensuring that the beamformers remain close to a good local optimum as the SNR increases.

Allowing several thousand iterations for convergence would likely be infeasible in practice. Hence achieving local sum-rate optimality is difficult in a distributed setting in which the channel and traffic conditions do not remain constant for a very long time. Realistically, the goal of the iterative updates could be to improve an initial set of beamformers within a limited number of iterations. From this point

[15]Also, Table II indicates that the weighted MSE algorithm may require fewer iterations to converge at moderate SNRs than the incremental SNR algorithm.

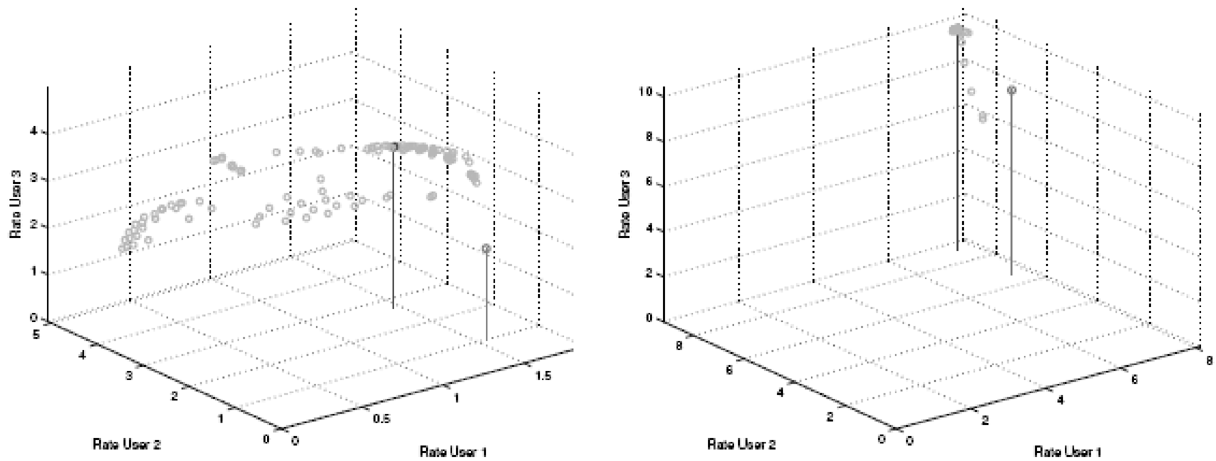


Fig. 6. Points in achievable rate region obtained by minimizing weighted sum MSE. Parameters are $K = 3$, $N = M = 2$, and SNR equal to $10\,\mathrm{dB}$ (left) and $30\,\mathrm{dB}$ (right).

of view, the results in Fig. 5 indicate that the algorithms considered here can deliver substantial gains with a modest number of iterations.

D. Rate Region

Fig. 6 shows achievable points in the rate region obtained by varying the MSE weights in the AW-MSE algorithm for $K = 3$ and $N = M = 2$. Two plots are shown for a particular channel realization with SNRs of $10\,\mathrm{dB}$ and $30\,\mathrm{dB}$. Two additional points are shown corresponding to the max-SINR and min-leakage algorithms. Those are connected to the bottom plane by solid lines, and essentially correspond to the two aligned solutions. The plots show that as the SNR increases, the points cluster around the aligned solution with the largest sum rate. This is to be expected since as the SNR becomes large, the rate region looks approximately like a box (with rounded edges), where the corner point corresponds to this aligned solution.

VI. CONCLUSIONS

We have compared properties and contrasted the performance of different algorithms for updating beamformers and receivers in a MIMO interference network, assuming each user transmits a single beam. The algorithms can be classified into two groups, depending on whether or not they support power control. The fixed-power algorithms (e. g., max-SINR and minimum-leakage) are observed to achieve the maximum DoFs at high SNRs when the corresponding number of users are activated. In general settings it may be necessary to combine such algorithms with a scheduler that limits the set of active users. For high SNRs, the number of active users should correspond to the maximum DoFs; however, determining the number of users that maximizes sum rate at moderate or low SNR depends on the direct- and cross-channel gains, and may require more information exchange in a distributed network. Although the max-SINR algorithm does not explicitly attempt to maximize sum-rate, our results show that it offers the best performance among fixed-power algorithms in terms of sum rate and convergence. Providing analytical justification for this observation remains an open and challenging problem.

The algorithms in the second group (including AW-MSE, interference pricing, and incremental SNR) are able to deactivate individual users, thus inherently performing user selection. In principle, they can therefore achieve the maximum DoFs when the system is overloaded, although the DoFs actually achieved depend on the system and algorithm parameters, and may be less than the maximum for particular channel realizations. These algorithms can be designed to maximize a sum utility objective, which enables proofs of convergence and provides additional flexibility in achieving different points in the rate region. However, this additional flexibility of being able to adjust the beam powers leads to slower and less reliable convergence. Among the algorithms considered here, the incremental SNR algorithm yielded the largest sum-rate in most cases, and exhibited reliable convergence. Again, providing analytical insight into these observations appears to be quite challenging.

It remains to compare these algorithms for the general case of multiple beams per user. While the max-SINR algorithm has been considered for more than a single beam [12], the performance is sensitive to the number of streams activated per user. A challenge is to design beamforming algorithms with power control and variable stream allocation that reliably converge to a close-to-optimal set of precoders even when the optimal stream allocation is not known.

APPENDIX

TRANSMITTER UPDATE IN THE INTERFERENCE PRICING ALGORITHM

We define the matrix $\boldsymbol{B}$ to contain the vectors $\sqrt{\pi_j}\boldsymbol{H}_{jk}^{\mathrm{H}}\boldsymbol{g}_j^*$ for all $j \neq k$ as columns and the scalar $c = \sum_{j\neq k}|\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kj}\boldsymbol{v}_j|^2 + \sigma^2$. With the pseudo-inverse $\boldsymbol{B}^+$ we distinguish between two cases:

1) $(\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^+)\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^* \neq \mathbf{0}$ (zero-forcing is possible):

   We must determine the unit-norm vector $\boldsymbol{w}$, such that $\boldsymbol{A}\boldsymbol{w} = \lambda\boldsymbol{w}$ with $\lambda > 0$ and

$$\boldsymbol{A} = \frac{1}{|\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kk}\boldsymbol{w}|^2 + c}\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kk} - \boldsymbol{B}\boldsymbol{B}^{\mathrm{H}}.$$

   Since $\boldsymbol{A}$ depends on $\boldsymbol{w}$, the solution cannot simply be found by performing an eigenvalue decomposition of $\boldsymbol{A}$. Instead, we perform a bisection line search: since $\|\boldsymbol{w}\|_2^2 = 1$, the value of $|\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kk}\boldsymbol{w}|^2$

must be in the interval $[0, \|\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*\|_2^2]$. We can now assume a fixed value of $|\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kk}\boldsymbol{w}|^2$ somewhere in this interval and compute the eigenvector corresponding to the unique positive eigenvalue of $\boldsymbol{A}$: if the value of $|\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kk}\boldsymbol{w}|^2$ resulting from the EVD of $\boldsymbol{A}$ is larger (smaller) than the fixed value of $|\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kk}\boldsymbol{w}|^2$ assumed for the computation of $\boldsymbol{A}$, then the fixed value was chosen too low (too high), cf. [40], and we have thus found a new lower (upper) bound for $|\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kk}\boldsymbol{w}|^2$. Once the correct vector $\boldsymbol{w}$ is found with the desired precision by continually reducing the interval for $|\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kk}\boldsymbol{w}|^2$, the updated beamformer is $\boldsymbol{v}_k^{\mathrm{new}} = \boldsymbol{w}$.

2) $(\mathbf{I} - \boldsymbol{B}\boldsymbol{B}^+)\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^* = \mathbf{0}$ (zero-forcing is not possible):

We compute

$$\zeta = \|\boldsymbol{B}^+\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*\|_2^2 - c$$

and distinguish between the following three cases:

a) $\zeta \leq 0$:

The updated beamformer is

$$\boldsymbol{v}_k^{\mathrm{new}} = \mathbf{0}.$$

b) $0 < \zeta \leq \|\boldsymbol{B}^+\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*\|_2^4 / \|(\boldsymbol{B}\boldsymbol{B}^{\mathrm{H}})^+\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*\|_2^2$:

The updated beamformer is

$$\boldsymbol{v}_k^{\mathrm{new}} = \frac{\sqrt{\zeta}}{\|\boldsymbol{B}^+\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*\|_2^2}(\boldsymbol{B}\boldsymbol{B}^{\mathrm{H}})^+\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*.$$

c) $\zeta > \|\boldsymbol{B}^+\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*\|_2^4 / \|(\boldsymbol{B}\boldsymbol{B}^{\mathrm{H}})^+\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*\|_2^2$:

As in case 1), we must determine the unit norm vector $\boldsymbol{w}$, such that $\boldsymbol{A}\boldsymbol{w} = \lambda\boldsymbol{w}$ with $\lambda > 0$ and

$$\boldsymbol{A} = \frac{1}{|\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kk}\boldsymbol{w}|^2 + c}\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kk} - \boldsymbol{B}\boldsymbol{B}^{\mathrm{H}}.$$

Again, we can use a bisection line search; in this case, the initial interval for the value of $|\boldsymbol{g}_k^{\mathrm{T}}\boldsymbol{H}_{kk}\boldsymbol{w}|^2$ is $[0, \min\{\zeta, \|\boldsymbol{H}_{kk}^{\mathrm{H}}\boldsymbol{g}_k^*\|_2^2\}]$. The updated beamformer is $\boldsymbol{v}_k^{\mathrm{new}} = \boldsymbol{w}$.

References

[1] D. Gesbert, S. Hanly, H. Huang, S. Shamai (Shitz), O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, 2010.

[2] V. R. Cadambe and S. A. Jafar, "Interference alignment and degrees of freedom of the K-user interference channel," *IEEE Trans. Inf. Theory*, vol. 54, no. 8, pp. 3425–3441, Aug. 2008.

[3] A. S. Motahari, S. Oveis-Gharan, M.-A. Maddah-Ali, and A. K. Khandani, "Real interference alignment: Exploiting the potential of single antenna systems," arXiv.org, Nov. 2009. [Online]. Available: http://arxiv.org/abs/0908.2282v2

[4] S. Sridharan, A. Jafarian, S. Vishwanath, S. A. Jafar, and S. Shamai (Shitz), "A layered lattice coding scheme for a class of three user Gaussian interference channels," in *Proc. 46th Annu. Allerton Conf. Commun., Control, Computing*, Monticello, IL, Sep. 2008, pp. 531–538.

[5] R. Etkin and E. Ordentlich, "On the degrees-of-freedom of the K-user Gaussian interference channel," arXiv.org, Jan. 2009. [Online]. Available: http://arxiv.org/abs/0901.1695v1

[6] C. Shi, D. A. Schmidt, R. A. Berry, M. L. Honig, and W. Utschick, "Distributed interference pricing for the MIMO interference channel," in *Proc. IEEE Int. Conf. Commun. 2009*, Dresden, Germany, Jun. 2009.

[7] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, Sep. 2010.

[8] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Utschick, "Distributed resource allocation schemes," *IEEE Signal Process. Mag.*, vol. 26, no. 5, pp. 53–63, Sep. 2009.

[9] D. A. Schmidt, W. Utschick, and M. L. Honig, "Beamforming techniques for single-beam MIMO interference networks," in *Proc. 48th Annu. Allerton Conf. Commun., Control, Computing*, Monticello, IL, Sep. 2010.

[10] C. M. Yetis, T. Gou, S. A. Jafar, and A. H. Kayran, "On feasibility of interference alignment in MIMO interference networks," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4771–4782, Sep. 2010.

[11] D. A. Schmidt, W. Utschick, and M. L. Honig, "Large system performance of interference alignment in single-beam MIMO networks," in *Proc. IEEE Global Commun. Conf. (GLOBECOM) 2010*, Miami, FL, Dec. 2010.

[12] K. Gomadam, V. R. Cadambe, and S. A. Jafar, "Approaching the capacity of wireless networks through distributed interference alignment," arXiv.org, Mar. 2008. [Online]. Available: http://arxiv.org/abs/0803.3816v1

[13] S. Stanczak, M. Kaliszand, and N. Bambos, "Admission control for power-controlled wireless networks under general interference functions," in *Proc. Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, Ca, Nov. 2008.

[14] J. Huang, R. A. Berry, and M. L. Honig, "Distributed interference compensation for wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 5, pp. 1074–1084, May 2006.

[15] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Utschick, "Minimum mean squared error interference alignment," in *Proc. 43rd Asilomar Conf. Signals, Syst., Comput.*, Pacific Grove, CA, Nov. 2009.

[16] S. S. Christensen, R. Agarwal, E. de Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[17] G. Bresler, D. Cartwright, and D. Tse, "Settling the feasibility of interference alignment for the MIMO interference channel: the symmetric square case," arXiv.org, Apr. 2011. [Online]. Available: http://arxiv.org/abs/1104.0888v1

[18] M. Razaviyayn, G. Lyubeznik, and Z.-Q. Luo, "On the degrees of freedom achievable through interference alignment in a MIMO interference channel," arXiv.org, Sep. 2011. [Online]. Available: http://arxiv.org/abs/1104.0992v2

[19] I. Santamaria, O. Gonzales, R. W. Heath Jr., and S. W. Peters, "Maximum sum-rate interference alignment algorithms for MIMO channels," in *Proc. IEEE Global Commun. Conf. (GLOBECOM) 2010*, Miami, FL, Dec. 2010.

[20] H. Sung, S.-H. Park, K.-J. Lee, and I. Lee, "Linear precoder designs for K-user interference channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 291–301, Jan. 2010.

[21] K. R. Kumar and F. Xue, "An iterative algorithm for joint signal and interference alignment," in *Proc. IEEE Int. Symp. Inform. Theory 2010*, Austin, TX, Jun. 2010, pp. 2293–2297.

[22] S. W. Peters and R. W. Heath, Jr., "Cooperative algorithms for MIMO interference channels," *IEEE Trans. Veh. Technol.*, vol. 60, no. 1, pp. 206–218, Jan. 2011.

[23] S. Ye and R. S. Blum, “Optimized signaling for MIMO interference systems with feedback,” *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.

[24] M. Razaviyayn, M. Sanjabi, and Z.-Q. Luo, “Linear transceiver design for interference alignment: Complexity and computation,” arXiv.org, Sep. 2010. [Online]. Available: http://arxiv.org/abs/1009.3481v1

[25] J. Papandriopoulos and J. S. Evans, “SCALE: A low-complexity distributed protocol for spectrum balancing in multiuser DSL networks,” *IEEE Trans. Inf. Theory*, vol. 55, no. 8, pp. 3711–3724, Aug. 2009.

[26] O. González and I. Santamaría, “Interference alignment in single-beam MIMO networks via homotopy continuation,” in *Proc. 36th Int. Conf. Acoust., Speech, Signal Process.*, Prague, Czech Republic, May 2011.

[27] L.-P. Qian, Y.-J. Zhang, and J. Huang, “MAPEL: Achieving global optimality for a non-convex wireless power control problem,” *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1553–1563, Mar. 2009.

[28] E. A. Jorswieck and E. G. Larsson, “Monotonic optimization framework for the two-user MISO interference channel,” *IEEE Trans. Commun.*, vol. 58, no. 7, pp. 2159–2168, Jul. 2010.

[29] G. Rajappan and M. L. Honig, “Signature sequence adaptation for DS-CDMA with multipath,” *IEEE J. Sel. Areas Commun.*, vol. 20, no. 2, pp. 384–395, Feb. 2002.

[30] M. F. Demirkol and M. A. Ingram, “Power-controlled capacity for interfering MIMO links,” in *Proc. 54th IEEE Veh. Technol. Conf. Fall 2001*, vol. 1, Atlantic City, NJ, Oct. 2001, pp. 187–191.

[31] ——, “Stream control in networks with interfering MIMO links,” in *Proc. IEEE Wireless Commun., Netw. Conf. 2003*, New Orleans, LA, Mar. 2003.

[32] G. Arslan, M. F. Demirkol, and Y. Song, “Equilibrium efficiency improvement in MIMO interference systems: A decentralized stream control approach,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 8, pp. 2984–2993, Aug. 2007.

[33] G. Scutari, D. P. Palomar, and S. Barbarossa, “The MIMO iterative waterfilling algorithm,” *IEEE Trans. Signal Process.*, vol. 57, no. 5, pp. 1917–1935, May 2009.

[34] C. Shi, R. Berry, and M. L. Honig, “Adaptive beamforming in interference networks via bidirectional training,” in *Conf. on Inf. Sciences and Systems (CISS)*, Princeton, NJ, Mar. 2010.

[35] R. Zakhour and D. Gesbert, “Coordination on the MISO interference channel using the virtual SINR framework,” in *Proc. Int. ITG Workshop Smart Antennas 2009*, Berlin, Germany, Feb. 2009.

[36] J. Park, Y. Sung, and H. V. Poor, “On beamformer design for multiuser MIMO interference channels,” arXiv.org, Nov. 2010. [Online]. Available: http://arxiv.org/abs/1011.6121v1

[37] H. Shen, B. Li, M. Tao, and X. Wang, “MSE-based transceiver designs for the MIMO interference channel,” *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3480–3489, Nov. 2010.

[38] W. M. Jang, B. R. Vojčić, and R. Pickholtz, “Joint transmitter-receiver optimization in synchronous multiuser communications over multipath channels,” *IEEE Trans. Commun.*, vol. 46, no. 2, pp. 269–278, Feb. 1998.

[39] S. Serbetli and A. Yener, “Transceiver optimization for multiuser MIMO systems,” *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 214–226, Jan. 2004.

[40] D. A. Schmidt, “Communications in multi-antenna interference networks,” Ph.D. dissertation, Technische Universität München, Dec. 2011. [Online]. Available: http://mediatum.ub.tum.de/node?id=1079376

[41] D. A. Schmidt, A. Gründinger, W. Utschick, and M. L. Honig, “Distributed precoder optimization for interfering MISO channels,” in *Proc. Int. ITG Workshop Smart Antennas 2008*, Darmstadt, Germany, Feb. 2008.

[42] C. Shi, R. A. Berry, and M. L. Honig, “Monotonic convergence of distributed interference pricing in wireless networks,” in *Proc. IEEE Int. Symp. Inform. Theory 2009*, Seoul, Korea, Jun. 2009, pp. 1619–1623.

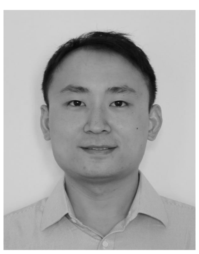
[43] R. Hunger, “Floating point operations in matrix-vector calculus,” Technische Universität München, Tech. Rep., Sep. 2007. [Online]. Available: http://mediatum2.ub.tum.de/doc/625604/625604.pdf

[44] S.-H. Park, H. Park, Y.-D. Kim, and I. Lee, “Regularized interference alignment based on weighted sum-MSE criterion for MIMO interference channels,” in *Proc. IEEE Int. Conf. Commun. 2010*, Cape Town, South Africa, May 2010.

David A. Schmidt received the Dipl.-Ing. degree in 2005 and the Dr.-Ing. degree in 2011, both from Technische Universität München (TUM), Munich, Germany.

From 2005 until 2011 he was a member of the scientific staff at the Associate Institute for Signal Processing at TUM, where he worked on signal processing techniques for MIMO interference channels; in the spring of 2009 he was a visiting researcher at Northwestern University, Evanston, IL. In 2011 he joined the test and measurement division of Rohde & Schwarz in Munich, Germany, where he currently develops hardware implementations of signal processing algorithms.

Dr. Schmidt is the recipient of the IEEE GLOBECOM 2010 GOLD Best Paper Award.



Changxin Shi received the B.E. degree in Electronic Engineering from Shanghai Jiao Tong University, China, in 2006, and the M.S. and Ph.D. degrees in Electrical Engineering and Computer Science from Northwestern University in 2008 and 2011, respectively. In 2012, he joined Barclays Capital in New York, NY, where he works as a quantitative analyst. His research interests include wireless communications with focus on resource allocation and dynamic spectrum sharing.

He is the recipient of 2012 IEEE Wireless Communications and Networking Conference (WCNC) Best Paper Award, a 2006 Walter P. Murphy Fellowship at Northwestern University, and a First-Class Excellent Academic Scholarship at Shanghai Jiao Tong University in both 2003 and 2005.

Randall A. Berry received the B.S. degree in Electrical Engineering from the University of Missouri-Rolla in 1993 and the M.S. and PhD degrees in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology in 1996 and 2000, respectively. In 2000, he joined Northwestern University, where he is currently a Professor in the Department of Electrical Engineering and Computer Science. In 1998 he was on the technical staff at MIT Lincoln Laboratory in the Advanced Networks Group. His primary research interests include wireless communication, network economics, data networks, and information theory.

Dr. Berry has served as an editor for the IEEE Transactions on Wireless Communications (2006-2009) and the IEEE Transactions on Information Theory (2009-2011). He has also been a guest editor for special issues of the IEEE Journal on Selected Topics in Signal Processing and the IEEE Transactions on Information Theory. He was a co-chair of the 2012 IEEE Communication Theory Workshop and a technical co-chair of 2010 IEEE ICC Wireless Networking Symposium. He is the recipient of a 2003 NSF CAREER award.

Michael L. Honig received the B.S. degree in electrical engineering from Stanford University in 1977, and the M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1978 and 1981, respectively. He subsequently joined Bell Laboratories in Holmdel, NJ, where he worked on local area networks and voiceband data transmission. In 1983 he joined the Systems Principles Research Division at Bellcore, where he worked on Digital Subscriber Lines and wireless communications. Since the Fall of 1994, he has been with Northwestern University where he is a Professor in the Electrical and Computer Engineering Department. He has held visiting scholar positions at the Naval Research Laboratory (San Diego), the University of California, Berkeley, the University of Sydney, Princeton University, and the Technical University of Munich. He has also worked as a free-lance trombonist.

Dr. Honig has served as an editor for the IEEE Transactions on Information Theory (1998-2000) and the IEEE Transactions on Communications (1990-1995), and as guest editor for the Journal on Selected Topics in Signal Processing, European Transactions on Telecommunications, and Wireless Personal Communications. He has also served as a member of the Digital Signal Processing Technical Committee for the IEEE Signal Processing Society, and as a member of the Board of Governors for the Information Theory Society (1997-2002). He is a Fellow of IEEE, the recipient of a Humboldt Research Award for Senior U.S. Scientists, and the co-recipient of the 2002 IEEE Communications Society and Information Theory Society Joint Paper Award and the 2010 IEEE Marconi Prize Paper Award.

Wolfgang Utschick (SM'06) was born on May 6, 1964. He completed several industrial education programs before he received the diploma and doctoral degrees, both with honors, in electrical engineering from the Technische Universität München, Germany (TUM), in 1993 and 1998, respectively. In this period he held a scholarship of the Bavarian Ministry of Education for exceptional students. From 1998 to 2002, he codirected the Signal Processing Group of the Institute of Circuit Theory and Signal Processing, TUM. Since 2000, he has been consulting in 3GPP standardization in the field of multielement antenna systems. In 2002, he has been appointed Professor at the TUM where he is Head of the Fachgebiet Methoden der Signalverarbeitung. He teaches courses on signal processing, stochastic processes and optimization theory in the field of digital communications. Dr. Utschick was awarded in 2006 for his excellent teaching records at TUM, and in 2007 received the ITG Award of the German Society for Information Technology (ITG). He is a senior member of the German VDE/ITG where he has been appointed in the Expert Committee for Information and System Theory in 2009. He is currently also serving as a Chairman of the national DFG Focus Program 'Communications in Interference Limited Networks' (COIN). Wolfgang Utschick is an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING and an editor and founder of the Springer Book Series Foundations in Signal Processing, Communications and Networking. Since 2010 he serves for the IEEE Signal Processing Society as an elected member of the Technical Committee of Signal Processing for Communications and Networking.