Gradient Projection Based Interference Alignment for the MIMO X Channel

Andreas Dotzler,* David Schmidt,* Guido Dietl,† and Wolfgang Utschick*

* Associate Institute for Signal Processing, Technische Universität München, 80290 Munich
E-Mail: {dotzler,dschmidt,utschick}@tum.de, Tel.: +49-89-289-28510, Fax: +49-89-289-28504

† DOCOMO Communications Laboratories Europe GmbH, Landsberger Str. 312, 80687 Munich
E-Mail: dietl@docomolab-euro.com, Tel.: +49-89-56824-245 Fax: +49-89-56824-301

***Abstract*—We regard a MIMO network with two transmitters and two receivers, in which each transmitter sends information to both of the receivers, a scenario known as the MIMO X channel. For such a system, the recently proposed interference alignment technique is proven to achieve the maximum degrees of freedom, which cannot be reached by conventional zero-forcing. In contrast to other scenarios, for the MIMO X channel algebraic expressions to obtain interference alignment can be easily found. Additionally, the set of parameters for possible alignments is a continuous manifold rather than a discrete set, which directly raises the question of how to find the best alignment when aiming at maximizing a utility of the transmission rates. Due to the non-convexity of the problem, finding the global optimally solution is numerically exhaustive and we are willing to accept a locally optimal solution. In this work we show an efficient parametrization of the problem which allows to apply a projected gradient approach that guarantees an aligned solution. In numerical simulations we show the superiority of our method compared to existing algorithms.**

I. Introduction

Interference management is a major challenge for the development of future wireless communication systems. Multiple antennas at the transmitter and receiver allow to utilize the additional spatial dimension to reduce or completely nullify the interference caused by transmission to other users. However, completely avoiding interference comes at the price of serving fewer users, as the degrees of freedom (DoF) in a wireless network are limited. The achievable DoF are given by the multiplicative increase of the sum-rate R, in the high power regime:

$$D = \lim_{P_{\text{tx}} \to \infty} \frac{R}{\log(P_{\text{tx}})}.$$

The concept of interference alignment (IA) is the main tool for achieving DoF higher than previously assumed [1], [2]. The two user MIMO X channel, is the smallest network for which IA can be applied and it has been shown that by time or frequency extensions of the channel IA achieves the highest number of DoF [2]. The upper bound for the symmetric MIMO X channel, where each receiver and each transmitter is equipped with N antennas, is $\frac{4N}{3}$, which can be achieved whenever N is a multiple of three, or the channel is extended to a multiple of three. Therefore, in this paper we focus on the case where $N = 3$, which directly extends to the case

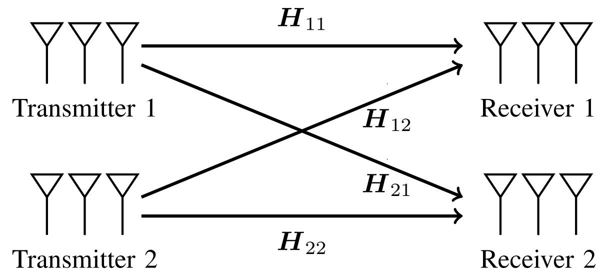

Fig. 1. The Three Antenna MIMO X Channel

where N is a multiple of three. Despite other setups where alignments can only be found by iterative algorithms, as for example [3], for the MIMO X channel there are algebraic expressions how to construct transmit and receive filters to align interference. However, the solution to the alignment problem is neither unique nor discrete, instead the set of alignments is a continuous manifold.

A scenario where the set of aligned solutions is a discrete set is the three user interference channel where each receiver and transmitter has two antennas and one data stream per transmitter receiver pair can be conveyed orthogonally. Starting from the conditions for interference free communication some simple calculus reveals that there are exactly two configurations where interference is aligned. Having fixed receive and transmit filters to one of the configurations it is clear that all transmitters use their complete power budget, so for infinite transmit power there exist exactly two relevant transmission strategies. Taking into account that for moderate transmit power one might refrain from strict alignment in order to improve the individual rates, the solution space becomes continuous and optimizing the transmission strategies is regarded in [4]. In the scenario regarded here, each transmitter sends information to both of the receivers and both streams share a power budget which can be freely allocated between them and for fixed directions of the filters rates can be continuously assigned by power allocation. Surprisingly there are infinitely many spatial configurations of the transmit and receive filters that align interference. In our work we optimize linear transmit and receive filters in order to maximize sum-rate. Finding good alignments was also considered for example in [1], [4], and [3], which apply alternating optimization, by either switching between the transmitters, or the original and reciprocal network.

After introducing the system model, we state a rule on how to select transmit and receive filters in order to obtain an aligned solution. We then formulate the sum-rate optimization problem where the constraints are chosen such that an aligned solution is guaranteed. We are able to eliminate the receive filters and two of the transmit filters and show how a projected gradient method can be used to compute a locally optimal solution. Further we introduce an alternative parametrization of the problem that greatly simplifies the projection step. The derivation of the gradient is given in detail and we present simulation results to illustrate the performance of our method.

II. System Model and Problem Statement

A. Notation

Throughout this paper, we use boldface upper and boldface lower case letters for matrices and vectors, respectively. $\mathbf{0}$ is a matrix where all elements are zero and $\boldsymbol{I}$ is and identity matrix where the size follows from the context. Operators $(\cdot)^{\mathrm{T}}$, $(\cdot)^{*}$, and $(\cdot)^{\mathrm{H}}$ denote the matrix transpose, conjugate, and conjugate transpose, respectively. We use $[\boldsymbol{M}]_{[k,:]}$ to select the k-th row, $[\boldsymbol{M}]_{[:,l]}$ to select the l-th column, and $[\boldsymbol{M}]_{[k,l]}$ to select the element in the k-th row and l-th column of a matrix. We use $\partial \boldsymbol{M}/\partial \xi$ as an abbreviation for elementwise derivation of the matrix elements with respect to ξ.

B. The MIMO X Channel

The MIMO X channel, illustrated in Figure 1, has two transmitters $t \in \mathcal{T}$ and two receivers $r \in \mathcal{R}$, where each transmitter is allowed to send information to both of the receivers. We denote the set of transmitter-receiver pairs as

$$\mathcal{K} = \{(rt) \in \mathcal{R} \times \mathcal{T}\}.$$

The DoF under conventional zero-forcing for the symmetric MIMO X channel with $N = 3$ are three, which can easily be achieved by activating only one point-to-point link in the system. In order to achieve DoF of four by interference alignment each transmitter sends one data stream to each receiver. The four scalar data symbols $s_{11}, s_{12}, s_{21}, s_{22}$ are filtered with the linear transmit filters $\boldsymbol{v}_{11}, \boldsymbol{v}_{12}, \boldsymbol{v}_{21}, \boldsymbol{v}_{22} \in \mathbb{C}^3$ and sent over the channels $\boldsymbol{H}_{11}, \boldsymbol{H}_{12}, \boldsymbol{H}_{21}, \boldsymbol{H}_{22} \in \mathbb{C}^{3\times 3}$, which we assume to have full rank. The received signals at the two users are

$$\begin{aligned}
\boldsymbol{y}_1 &= \underbrace{\boldsymbol{H}_{11}\boldsymbol{v}_{11}s_{11}}_{\text{intended signal}} + \underbrace{\boldsymbol{H}_{11}\boldsymbol{v}_{21}s_{21}}_{\text{interference}} + \\
&\quad + \underbrace{\boldsymbol{H}_{12}\boldsymbol{v}_{12}s_{12}}_{\text{intended signal}} + \underbrace{\boldsymbol{H}_{12}\boldsymbol{v}_{22}s_{22}}_{\text{interference}} + \boldsymbol{n}_1, \\
\boldsymbol{y}_2 &= \underbrace{\boldsymbol{H}_{21}\boldsymbol{v}_{11}s_{11}}_{\text{interference}} + \underbrace{\boldsymbol{H}_{21}\boldsymbol{v}_{21}s_{21}}_{\text{intended signal}} + \\
&\quad + \underbrace{\boldsymbol{H}_{22}\boldsymbol{v}_{12}s_{12}}_{\text{interference}} + \underbrace{\boldsymbol{H}_{22}\boldsymbol{v}_{22}s_{22}}_{\text{intended signal}} + \boldsymbol{n}_2,
\end{aligned}$$

where $\boldsymbol{n}_1, \boldsymbol{n}_2 \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \boldsymbol{I})$ represent additive white Gaussian noise. Assuming Gaussian signalling and treating interference as additional noise, the rates of each stream are given by

$$R_{rt} = \log\left(1 + \frac{\left|\boldsymbol{g}_{rt}^{\mathrm{H}}\boldsymbol{H}_{rt}\boldsymbol{v}_{rt}\right|^2}{\sum_{(ij)\in\mathcal{K}\setminus(rt)} \left|\boldsymbol{g}_{rt}^{\mathrm{H}}\boldsymbol{H}_{rj}\boldsymbol{v}_{ij}\right|^2 + \boldsymbol{g}_{rt}^{\mathrm{H}}\boldsymbol{g}_{rt}\sigma^2}\right),$$

where $\boldsymbol{g}_{rt}$ are the linear receive filters. The noise plus interference matrix for each stream is given by

$$\boldsymbol{X}_{rt} = \sum_{(ij)\in\mathcal{K}\setminus(rt)} \boldsymbol{H}_{rj}\boldsymbol{v}_{ij}\boldsymbol{v}_{ij}^{\mathrm{H}}\boldsymbol{H}_{rj}^{\mathrm{H}} + \sigma^2\boldsymbol{I}$$

Assuming the transmit filters are fixed, the optimal receive filters are given by

$$\boldsymbol{g}_{rt}^{\mathrm{H}} = \boldsymbol{v}_{rt}^{\mathrm{H}}\boldsymbol{H}_{rt}^{\mathrm{H}}\boldsymbol{X}_{rt}^{-1}$$

and the rates can be expressed as

$$R_{rt} = \log\left(1 + \boldsymbol{v}_{rt}^{\mathrm{H}}\boldsymbol{H}_{rt}^{\mathrm{H}}\boldsymbol{X}_{rt}^{-1}\boldsymbol{H}_{rt}\boldsymbol{v}_{rt}\right).$$

The sum-rate optimization problem, w.l.o.g. chosen as utility function, under transmit power constraints $P_{\mathrm{tx},1}, P_{\mathrm{tx},2}$, can then be stated as

$$\begin{aligned}
&\underset{\boldsymbol{v}_{rt},(rt)\in\mathcal{K}}{\text{maximize}} && \sum_{(rt)\in\mathcal{K}} R_{rt} \\
&\text{subject to} && ||\boldsymbol{v}_{1t}||_2^2 + ||\boldsymbol{v}_{2t}||_2^2 \leq P_{\mathrm{tx},t},\ \forall\, t \in \mathcal{T}
\end{aligned}$$

It is clear that for sufficiently high transmit power, the optimal solution to the optimization problem are transmit filters such that interference is aligned and the receive filters become the zero-forcing filters. However, due to the non-convexity of the problem, computing the optimal solution is numerically almost intractable and attempts to find a good solution by local methods, as for example the pricing methods used in [5], fail to find an aligned solution and one of the streams is shut off. In the following we present a reformulation of the problem, that allows for optimization while maintaining an aligned solution.

III. Gradient Projection Based Interference Alignment

A. Problem Statement

The condition for interference free transmission of four streams in the system is that the two interfering signals at each receiver are aligned in a one-dimensional space, which can be algebraically expressed as

$$\boldsymbol{H}_{11}\boldsymbol{v}_{21} = \lambda_1\boldsymbol{H}_{12}\boldsymbol{v}_{22}, \quad \boldsymbol{H}_{22}\boldsymbol{v}_{12} = \lambda_2\boldsymbol{H}_{21}\boldsymbol{v}_{11},$$

which directly allows us to eliminate two of the transmit filters,

$$\boldsymbol{v}_{21} = \lambda_1\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{v}_{22}, \quad \boldsymbol{v}_{12} = \lambda_2\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{v}_{11}. \tag{1}$$

The zero-forcing conditions for the receive filters are

$$\begin{aligned}
\boldsymbol{g}_{11}^{\mathrm{H}}\boldsymbol{H}_{12}\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{v}_{11} &= 0, & \boldsymbol{g}_{11}^{\mathrm{H}}\boldsymbol{H}_{12}\boldsymbol{v}_{22} &= 0,\\
\boldsymbol{g}_{12}^{\mathrm{H}}\boldsymbol{H}_{11}\boldsymbol{v}_{11} &= 0, & \boldsymbol{g}_{12}^{\mathrm{H}}\boldsymbol{H}_{12}\boldsymbol{v}_{22} &= 0,\\
\boldsymbol{g}_{21}^{\mathrm{H}}\boldsymbol{H}_{22}\boldsymbol{v}_{22} &= 0, & \boldsymbol{g}_{21}^{\mathrm{H}}\boldsymbol{H}_{21}\boldsymbol{v}_{11} &= 0,\\
\boldsymbol{g}_{22}^{\mathrm{H}}\boldsymbol{H}_{21}\boldsymbol{v}_{11} &= 0, & \boldsymbol{g}_{22}^{\mathrm{H}}\boldsymbol{H}_{21}\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{v}_{22} &= 0,
\end{aligned}$$

which we generalize as

$$\boldsymbol{g}_{rt}^{\mathrm{H}}\boldsymbol{A}_{rt}\boldsymbol{v}_{11} = 0,\ \boldsymbol{g}_{rt}^{\mathrm{H}}\boldsymbol{B}_{rt}\boldsymbol{v}_{22} = 0\ \forall\ (rt) \in \mathcal{K},$$

and notice that for given transmit filters the receive filters are constrained to a one-dimensional subspace. For fixed transmit filters we search for the normalized receive filter that maximizes the rate of the datastream under the constraint to nullify all interference, which is the solution of the following optimization problem:

$$\begin{aligned}
\underset{\boldsymbol{g}_{rt}}{\text{maximize}}\quad & \boldsymbol{g}_{rt}^{\mathrm{H}}\boldsymbol{H}_{rt}\boldsymbol{v}_{rt}\\
\text{subject to}\quad & \boldsymbol{g}_{rt}^{\mathrm{H}}\boldsymbol{A}_{rt}\boldsymbol{v}_{11} = 0,\\
& \boldsymbol{g}_{rt}^{\mathrm{H}}\boldsymbol{B}_{rt}\boldsymbol{v}_{22} = 0,\\
& ||\boldsymbol{g}_{rt}||_2 = 1.
\end{aligned} \tag{2}$$

By introducing a projection matrix

$$\boldsymbol{P}_{rt} = \boldsymbol{I} - \boldsymbol{C}_{rt}\left(\boldsymbol{C}_{rt}^{\mathrm{H}}\boldsymbol{C}_{rt}\right)^{-1}\boldsymbol{C}_{rt}^{\mathrm{H}}, \tag{3}$$

where

$$\boldsymbol{C}_{rt} = \left[\boldsymbol{A}_{rt}\boldsymbol{v}_{11}\boldsymbol{B}_{rt}\boldsymbol{v}_{22}\right], \tag{4}$$

it is easy to see that

$$\boldsymbol{g}_{rt}^{\mathrm{H}} = \frac{\boldsymbol{v}_{rt}^{\mathrm{H}}\boldsymbol{H}_{rt}^{\mathrm{H}}\boldsymbol{P}_{rt}}{\left\|\boldsymbol{v}_{rt}^{\mathrm{H}}\boldsymbol{H}_{rt}^{\mathrm{H}}\boldsymbol{P}_{rt}\right\|_2}$$

is a maximizer of the optimization problem. This way we eliminate the receive filters and the rates can be expressed as

$$\begin{aligned}
R_{11} &= \log\left(1 + \tfrac{1}{\sigma^2}\boldsymbol{v}_{11}^{\mathrm{H}}\boldsymbol{H}_{11}^{\mathrm{H}}\boldsymbol{P}_{11}\boldsymbol{H}_{11}\boldsymbol{v}_{11}\right)\\
R_{12} &= \log\left(1 + \tfrac{\lambda_2^2}{\sigma^2}\boldsymbol{v}_{11}^{\mathrm{H}}\boldsymbol{H}_{21}^{\mathrm{H}}\boldsymbol{H}_{22}^{-\mathrm{H}}\boldsymbol{H}_{12}^{\mathrm{H}}\boldsymbol{P}_{12}\boldsymbol{H}_{12}\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{v}_{11}\right)\\
R_{21} &= \log\left(1 + \tfrac{\lambda_1^2}{\sigma^2}\boldsymbol{v}_{22}^{\mathrm{H}}\boldsymbol{H}_{12}^{\mathrm{H}}\boldsymbol{H}_{11}^{-\mathrm{H}}\boldsymbol{H}_{21}^{\mathrm{H}}\boldsymbol{P}_{21}\boldsymbol{H}_{21}\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{v}_{22}\right)\\
R_{22} &= \log\left(1 + \tfrac{1}{\sigma^2}\boldsymbol{v}_{22}^{\mathrm{H}}\boldsymbol{H}_{22}^{\mathrm{H}}\boldsymbol{P}_{22}\boldsymbol{H}_{22}\boldsymbol{v}_{22}\right)
\end{aligned}$$

where we substituted $\boldsymbol{v}_{12}$ and $\boldsymbol{v}_{21}$ according to (1). The sum-rate optimization problem becomes

$$\begin{aligned}
\underset{\substack{\boldsymbol{v}_{11},\boldsymbol{v}_{22},\\ \lambda_1,\lambda_2}}{\text{maximize}}\quad & \sum_{(rt)\in\mathcal{K}} R_{rt}\\
\text{subject to}\quad & \|\boldsymbol{v}_{11}\|_2^2 + \left\|\lambda_1\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{v}_{22}\right\|_2^2 \le P_{\mathrm{tx},1},\\
& \|\boldsymbol{v}_{22}\|_2^2 + \left\|\lambda_2\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{v}_{11}\right\|_2^2 \le P_{\mathrm{tx},2}.
\end{aligned}$$

We would like to remark that we did not impose the necessary constraint that $\boldsymbol{v}_{11}$ and $\boldsymbol{v}_{21}$ as well as $\boldsymbol{v}_{12}$ and $\boldsymbol{v}_{22}$ must not be collinear. For channel coefficients drawn from a continuous distribution and random choices of $\boldsymbol{v}_{11}$ and $\boldsymbol{v}_{22}$, this constraint is fulfilled with probability one. This argument is, however, too weak for the algorithm we suggest in the following, where $\boldsymbol{v}_{11}$ and $\boldsymbol{v}_{22}$ are chosen by sequential projected gradient updates. However, in numerical simulations we did not encounter any cases where the two transmit filters at one transmitter are collinear and therefore it seams legitimate not to explicitly enforce this constraint.

B. Projected Gradient

Despite having a nice formulation of the problem, which guarantees to have an aligned solution for any choice of the parameters

$$\boldsymbol{d} = \left[\boldsymbol{v}_{11}^{\mathrm{T}}, \boldsymbol{v}_{22}^{\mathrm{T}}, \lambda_1, \lambda_2\right]^{\mathrm{T}},$$

the problem is still non-convex and therefore computing the global optimizers is out of reach. We use a projected gradient method, which uses an arbitrary feasible $\boldsymbol{d}^{(i)}$ and iteratively calculates a complex gradient of the real valued sum-rate function

$$\nabla\boldsymbol{d}^{(i)} = 2\cdot\sum_{(rt)\in\mathcal{K}}\begin{bmatrix}\frac{\partial R_{rt}^{(i)}}{\partial\boldsymbol{v}_{11}^*}\\ \frac{\partial R_{rt}^{(i)}}{\partial\boldsymbol{v}_{22}^*}\\ \frac{\partial R_{rt}^{(i)}}{\partial\lambda_1^*}\\ \frac{\partial R_{rt}^{(i)}}{\partial\lambda_2^*}\end{bmatrix},$$

and makes a step $\alpha^{(i)}\nabla\boldsymbol{d}^{(i)}$ into the direction of the gradient, where $\alpha^{(i)}$ is used to control the step size. However, the newly obtained parameters might be infeasible and therefore they are projected onto the feasible set

$$\boldsymbol{d}^{(i+1)} = \mathcal{P}\left(\boldsymbol{d}^{(i)} + \alpha^{(i)}\nabla\boldsymbol{d}^{(i)}\right).$$

The projected values are chosen from the feasible set such that the Euclidean norm of the distance to the newly obtained variables

$$\begin{aligned}
&\left\|\boldsymbol{d}^{(i+1)} - \left(\boldsymbol{d}^{(i)} + \alpha^{(i)}\nabla\boldsymbol{d}^{(i)}\right)\right\|_2 =\\
&= \left(\left\|\boldsymbol{v}_{11}^{(i+1)} - \left(\boldsymbol{v}_{11}^{(i)} + 2\alpha^{(i)}\sum_{(rt)\in\mathcal{K}}\frac{\partial R_{rt}^{(i)}}{\partial\boldsymbol{v}_{11}^*}\right)\right\|_2^2 +\right.\\
&\quad + \left\|\boldsymbol{v}_{22}^{(i+1)} - \left(\boldsymbol{v}_{22}^{(i)} + 2\alpha^{(i)}\sum_{(rt)\in\mathcal{K}}\frac{\partial R_{rt}^{(i)}}{\partial\boldsymbol{v}_{22}^*}\right)\right\|_2^2 +\\
&\quad + \left\|\lambda_1^{(i+1)} - \left(\lambda_1^{(i)} + 2\alpha^{(i)}\sum_{(rt)\in\mathcal{K}}\frac{\partial R_{rt}^{(i)}}{\partial\lambda_1^*}\right)\right\|_2^2 +\\
&\quad \left. + \left\|\lambda_2^{(i+1)} - \left(\lambda_2^{(i)} + 2\alpha^{(i)}\sum_{(rt)\in\mathcal{K}}\frac{\partial R_{rt}^{(i)}}{\partial\lambda_2^*}\right)\right\|_2^2\right)^{\frac{1}{2}}
\end{aligned} \tag{5}$$

is minimized and can be computed as the solution of the following optimization problem:

$$\begin{aligned}
\underset{\substack{\boldsymbol{v}_{11}^{(i)},\boldsymbol{v}_{22}^{(i)},\\ \lambda_1^{(i)},\lambda_2^{(i)}}}{\text{maximize}} \quad & \left\|\boldsymbol{d}^{(i+1)} - \left(\boldsymbol{d}^{(i)} + \alpha^{(i)}\nabla\boldsymbol{d}^{(i)}\right)\right\|_2^2 \\
\text{subject to} \quad & \left\|\boldsymbol{v}_{11}^{(i+1)}\right\|_2^2 + \left\|\lambda_1^{(i+1)}\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{v}_{22}^{(i+1)}\right\|_2^2 \leq P_{\text{tx},1} \\
& \left\|\boldsymbol{v}_{22}^{(i+1)}\right\|_2^2 + \left\|\lambda_2^{(i+1)}\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{v}_{11}^{(i+1)}\right\|_2^2 \leq P_{\text{tx},2}
\end{aligned}$$

which is convex and allows for efficient solution methods. In case $\alpha^{(i)}$ is selected such that $\alpha^{(i)}\nabla\boldsymbol{d}^{(i)}$ is an increasing direction and the norm of the projected gradient, given in (5), is zero, the points $\boldsymbol{v}_{11}^{(i+1)}, \boldsymbol{v}_{22}^{(i+1)}, \lambda_1^{(i+1)}, \lambda_2^{(i+1)}$ fulfill the first-order optimality conditions of the global maximizers and we accept them as locally optimal solution. Simulations show that the algorithm usually converges within very few iterations.

C. Reformulation of the Problem

To reduce complexity we derive a reformulation of the problem, so that the projection step is drastically simplified, but this comes at the price of a more complicated gradient. As a first step we decouple the power and spatial allocation by introducing variables P_{rt} corresponding to the power allocation per datastream:

$$\begin{aligned}
\boldsymbol{v}_{11} &= \frac{\boldsymbol{u}_{11}}{\|\boldsymbol{u}_{11}\|_2}\sqrt{P_{11}} \\
\boldsymbol{v}_{12} &= \frac{\lambda_2\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{u}_{11}}{\left\|\lambda_2\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{u}_{11}\right\|_2}\sqrt{P_{12}} = \frac{\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{u}_{11}}{\left\|\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{u}_{11}\right\|_2}\sqrt{P_{12}} \\
\boldsymbol{v}_{21} &= \frac{\lambda_1\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{u}_{22}}{\left\|\lambda_1\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{u}_{22}\right\|_2}\sqrt{P_{21}} = \frac{\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{u}_{22}}{\left\|\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{u}_{22}\right\|_2}\sqrt{P_{21}} \\
\boldsymbol{v}_{22} &= \frac{\boldsymbol{u}_{22}}{\|\boldsymbol{u}_{22}\|_2}\sqrt{P_{22}}.
\end{aligned}$$

The expressions for the rates are now

$$\begin{aligned}
R_{11} &= \log\left(1 + \frac{P_{11}}{\sigma^2}\frac{\boldsymbol{u}_{11}^{\mathrm{H}}\boldsymbol{H}_{11}^{\mathrm{H}}\boldsymbol{P}_{11}\boldsymbol{H}_{11}\boldsymbol{u}_{11}}{\|\boldsymbol{u}_{11}\|_2^2}\right) \\
R_{12} &= \log\left(1 + \frac{P_{12}}{\sigma^2}\frac{\boldsymbol{u}_{11}^{\mathrm{H}}\boldsymbol{H}_{21}^{\mathrm{H}}\boldsymbol{H}_{22}^{-\mathrm{H}}\boldsymbol{H}_{12}^{\mathrm{H}}\boldsymbol{P}_{12}\boldsymbol{H}_{12}\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{u}_{11}}{\left\|\boldsymbol{H}_{22}^{-1}\boldsymbol{H}_{21}\boldsymbol{u}_{11}\right\|_2^2}\right) \\
R_{21} &= \log\left(1 + \frac{P_{21}}{\sigma^2}\frac{\boldsymbol{u}_{22}^{\mathrm{H}}\boldsymbol{H}_{12}^{\mathrm{H}}\boldsymbol{H}_{11}^{-\mathrm{H}}\boldsymbol{H}_{21}^{\mathrm{H}}\boldsymbol{P}_{21}\boldsymbol{H}_{21}\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{u}_{22}}{\left\|\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{u}_{22}\right\|_2^2}\right) \\
R_{22} &= \log\left(1 + \frac{P_{22}}{\sigma^2}\frac{\boldsymbol{u}_{22}^{\mathrm{H}}\boldsymbol{H}_{22}^{\mathrm{H}}\boldsymbol{P}_{22}\boldsymbol{H}_{22}\boldsymbol{u}_{22}}{\|\boldsymbol{u}_{22}\|_2^2}\right).
\end{aligned} \tag{6}$$

The rate expressions do not depend on the norm of $\boldsymbol{u}_{11}$ and $\boldsymbol{u}_{22}$ but only on their direction, therefore they can be arbitrarily scaled and the sum-rate optimization problem is

$$\begin{aligned}
\underset{\substack{\boldsymbol{u}_{11},\boldsymbol{u}_{22},\\ P_{rt},(rt)\in\mathcal{K}}}{\text{maximize}} \quad & \sum_{(rt)\in\mathcal{K}} R_{rt} \\
\text{subject to} \quad & P_{1t} + P_{2t} \leq P_{\text{tx},t}, \forall\, t\in\mathcal{T}.
\end{aligned}$$

The gradient is now given by

$$\nabla\boldsymbol{d}^{(i)} = \sum_{(rt)\in\mathcal{K}} \begin{bmatrix} 2\frac{\partial R_{rt}^{(i)}}{\partial\boldsymbol{u}_{11}^*} \\ 2\frac{\partial R_{rt}^{(i)}}{\partial\boldsymbol{u}_{22}^*} \\ \frac{\partial R_{rt}^{(i)}}{\partial P_{11}} \\ \vdots \\ \frac{\partial R_{rt}^{(i)}}{\partial P_{22}} \end{bmatrix},$$

Clearly, it is possible to eliminate two of the power allocation parameters by adjusting the relative power allocation between the two streams per transmitter directly. This however does not decrease complexity and in order to allow for future extensions, see Section III-E, we prefer to have one parameter per stream.

As we will see later the gradient with respect to the power allocation (11) is always non-negative, which matches the intuition that increasing the power for an orthogonal stream is always beneficial, and the projection onto the feasible set is

$$\mathcal{P}\left(P_{rt}'\right) = P_{\text{tx},t}\left(\max\{P_{\text{tx},t}, \sum_{r\in\mathcal{R}} P_{1t}'\}\right)^{-1} P_{rt}',$$

where

$$P_{rt}' = P_{rt}^{(i)} + \alpha^{(i)}\sum_{(yz)\in\mathcal{K}}\frac{\partial R_{rt}^{(i)}}{\partial P_{yz}}.$$

For an implementation, it is convenient to rescale $\boldsymbol{u}_{11}$ and $\boldsymbol{u}_{22}$ to have norm one after every step, in order to avoid numerical problems.

D. Derivation of the Gradient

The prototype for the rate expressions in (6) is

$$R_{rt} = \log\left(1 + \frac{P_{rt}}{\sigma^2}\frac{\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}_{rt}^{\mathrm{H}}\boldsymbol{H}_{rt}^{\mathrm{H}}\boldsymbol{P}_{rt}\boldsymbol{H}_{rt}\boldsymbol{D}_{rt}\boldsymbol{u}_{ii}}{\|\boldsymbol{D}_{rt}\boldsymbol{u}_{ii}\|_2^2}\right),$$

for which we now derive $\partial R_{rt}/\boldsymbol{u}_{ii}^*, \partial R_{rt}/\partial\boldsymbol{u}_{jj}^*$, and $\partial R_{rt}/\partial P_{yz}$, where we treat $\boldsymbol{u}_{ii}$ and $\boldsymbol{u}_{jj}$ as fixed variables. For the sake of simpler notation we drop the index rt and start with

$$\frac{\partial R_{rt}}{\partial\boldsymbol{u}_{ii}^*} = \frac{1}{1+\frac{P}{\sigma^2}\gamma}\frac{P}{\sigma^2}\frac{\partial\gamma}{\partial\boldsymbol{u}_{ii}^*},$$

where

$$\gamma = \frac{\mu}{\nu} = \frac{\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}}{\|\boldsymbol{D}\boldsymbol{u}_{ii}\|_2^2}. \tag{7}$$

By applying the product rule we compute

$$\frac{\partial\gamma}{\partial\boldsymbol{u}_{ii}^*}=\frac{\nu\frac{\partial\mu}{\partial\boldsymbol{u}_{ii}^*}-\frac{\partial\nu}{\partial\boldsymbol{u}_{ii}^*}\mu}{\nu^2}=\frac{\|\boldsymbol{D}_{rt}\boldsymbol{u}_{ii}\|_2^2\frac{\partial\mu}{\partial\boldsymbol{u}_{ii}^*}-\frac{\partial\nu}{\partial\boldsymbol{u}_{ii}^*}\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}}{\|\boldsymbol{D}_{rt}\boldsymbol{u}_{ii}\|_2^4}=\frac{\|\boldsymbol{D}_{rt}\boldsymbol{u}_{ii}\|_2^2\frac{\partial\mu}{\partial\boldsymbol{u}_{ii}^*}-\boldsymbol{D}^{\mathrm{H}}\boldsymbol{D}\boldsymbol{u}_{ii}\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}}{\|\boldsymbol{D}_{rt}\boldsymbol{u}_{ii}\|_2^4}.$$

As the projector $\boldsymbol{P}$, defined in (3), itself depends on $\boldsymbol{u}_{ii}^*$, we compute $\partial\mu/\partial\boldsymbol{u}_{ii}^*$ elementwise for all entries $x=1,2,3$:

$$\frac{\partial\mu}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}=\frac{\partial\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}=\frac{\partial\sum_{k=1}^{3}\sum_{l=1}^{3}\left[\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right]_{[:,k]}[\boldsymbol{P}]_{[k,l]}[\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}]_{[l,:]}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}=\sum_{k=1}^{3}\sum_{l=1}^{3}\frac{\partial\left[\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right]_{[:,k]}[\boldsymbol{P}]_{[k,l]}[\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}]_{[l,:]}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}. \quad (8)$$

We notice that $[\boldsymbol{P}]_{[k,l]}$ always depends on $[\boldsymbol{u}_{ii}^*]_{[x,:]}$, however $\left[\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right]_{[:,k]}$ only when $k=x$, and therefore

$$\frac{\partial\left[\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right]_{[:,k]}[\boldsymbol{P}]_{[k,l]}[\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}]_{[l,:]}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}=\begin{cases}\left[\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right]_{[:,k]}\left[\frac{\partial\boldsymbol{P}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}\right]_{[k,l]}[\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}]_{[l,:]} & \text{if } k\neq x\\ \frac{\partial\left[\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right]_{[:,k]}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}[\boldsymbol{P}]_{[k,l]}[\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}]_{[l,:]}+\left[\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right]_{[:,k]}\left[\frac{\partial\boldsymbol{P}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}\right]_{[k,l]}[\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}]_{[l,:]} & \text{if } k=x\end{cases} \quad (9)$$

Plugging this result into (8), we obtain

$$\frac{\partial\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}=\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\frac{\partial\boldsymbol{P}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}+\frac{\partial\left[\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right]_{[:,x]}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}[\boldsymbol{P}]_{[x,:]}\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}=\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\frac{\partial\boldsymbol{P}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}+\left[\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\right]_{[x,x]}[\boldsymbol{P}]_{[x,:]}\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}$$

It remains to compute $\partial\boldsymbol{P}/\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}$ and in a first step we state

$$\frac{\partial\boldsymbol{P}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}=\frac{\partial\boldsymbol{I}-\boldsymbol{C}\left(\boldsymbol{C}^{\mathrm{H}}\boldsymbol{C}\right)^{-1}\boldsymbol{C}^{\mathrm{H}}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}=-\frac{\partial\boldsymbol{C}\left(\boldsymbol{C}^{\mathrm{H}}\boldsymbol{C}\right)^{-1}\boldsymbol{C}^{\mathrm{H}}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}=-\boldsymbol{C}\left(\boldsymbol{C}^{\mathrm{H}}\boldsymbol{C}\right)^{-1}\frac{\partial\boldsymbol{C}^{\mathrm{H}}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}-\frac{\partial\boldsymbol{C}\left(\boldsymbol{C}^{\mathrm{H}}\boldsymbol{C}\right)^{-1}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}\boldsymbol{C}^{\mathrm{H}} \quad (10)$$

where we used the definition of $\boldsymbol{C}$ given in (3). Knowing that

$$\frac{\partial\boldsymbol{M}^{-1}}{\partial\xi}=-\boldsymbol{M}^{-1}\frac{\partial\boldsymbol{M}^{-1}}{\partial\xi}\boldsymbol{M}^{-1},$$

we continue by stating

$$\frac{\partial\boldsymbol{P}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}=-\boldsymbol{C}\left(\boldsymbol{C}^{\mathrm{H}}\boldsymbol{C}\right)^{-1}\frac{\partial\boldsymbol{C}^{\mathrm{H}}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}+\boldsymbol{C}\left(\boldsymbol{C}^{\mathrm{H}}\boldsymbol{C}\right)^{-1}\frac{\partial\boldsymbol{C}^{\mathrm{H}}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}\boldsymbol{C}\left(\boldsymbol{C}^{\mathrm{H}}\boldsymbol{C}\right)^{-1}\boldsymbol{C}^{\mathrm{H}}=-\boldsymbol{C}\left(\boldsymbol{C}^{\mathrm{H}}\boldsymbol{C}\right)^{-1}\frac{\partial\boldsymbol{C}^{\mathrm{H}}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}\left(\boldsymbol{I}-\boldsymbol{C}\left(\boldsymbol{C}^{\mathrm{H}}\boldsymbol{C}\right)^{-1}\boldsymbol{C}^{\mathrm{H}}\right)=-\boldsymbol{C}\left(\boldsymbol{C}^{\mathrm{H}}\boldsymbol{C}\right)^{-1}\frac{\partial\boldsymbol{C}^{\mathrm{H}}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}\boldsymbol{P}.$$

The derivative of $\boldsymbol{C}^{\mathrm{H}}$, defined in (4), with respect to $[\boldsymbol{u}_{ii}^*]_{[x,:]}$ depends on whether i is 1 or 2 and is given by

$$\frac{\partial\boldsymbol{C}^{\mathrm{H}}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}=\frac{\partial\begin{bmatrix}\boldsymbol{u}_{11}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\\ \boldsymbol{u}_{22}^{\mathrm{H}}\boldsymbol{B}^{\mathrm{H}}\end{bmatrix}}{\partial[\boldsymbol{u}_{ii}^*]_{[x,:]}}\begin{cases}\begin{bmatrix}[\boldsymbol{A}]_{[x,:]}\\ \boldsymbol{0}\end{bmatrix} & \text{if } i=1,\\ \begin{bmatrix}\boldsymbol{0}\\ [\boldsymbol{B}]_{[x,:]}\end{bmatrix} & \text{if } i=2.\end{cases}$$

In contrast to the rather long derivation of $\partial R_{rt}/\partial\boldsymbol{u}_{ii}^*$, calculating $\partial R_{rt}/\partial\boldsymbol{u}_{jj}^*$ and $\partial R_{rt}/\partial P_{yz}$ is straightforward:

$$\frac{\partial R_{rt}}{\partial\boldsymbol{u}_{jj}^*}=\frac{1}{1+\frac{P}{\sigma^2}\gamma}\frac{P}{\sigma^2}\frac{\partial\gamma}{\partial\boldsymbol{u}_{jj}^*},$$

where γ is defined in (7). In a next step we can see that

$$\frac{\partial\gamma}{\partial\left[\boldsymbol{u}_{jj}^*\right]_{[x,:]}}=\frac{1}{\|\boldsymbol{D}_{rt}\boldsymbol{u}_{ii}\|_2^2}\frac{\partial\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\boldsymbol{P}\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii}}{\partial\left[\boldsymbol{u}_{jj}^*\right]_{[x,:]}}=\boldsymbol{u}_{ii}^{\mathrm{H}}\boldsymbol{D}^{\mathrm{H}}\boldsymbol{H}^{\mathrm{H}}\frac{\partial\boldsymbol{P}}{\partial\left[\boldsymbol{u}_{jj}^*\right]_{[x,:]}}\boldsymbol{H}\boldsymbol{D}\boldsymbol{u}_{ii},$$

and $\partial\boldsymbol{P}/\partial\left[\boldsymbol{u}_{jj}^*\right]_{[x,:]}$ is given in (10). As all streams are completely orthogonal, the rates only depend on their own power allocation and therefore:

$$\frac{\partial R_{rt}}{\partial P_{yz}}=\begin{cases}\dfrac{\frac{\gamma}{\sigma^2}}{1+\frac{P}{\sigma^2}\gamma} & \text{if } y=r \text{ and } z=t,\\ 0 & \text{otherwise.}\end{cases} \quad (11)$$

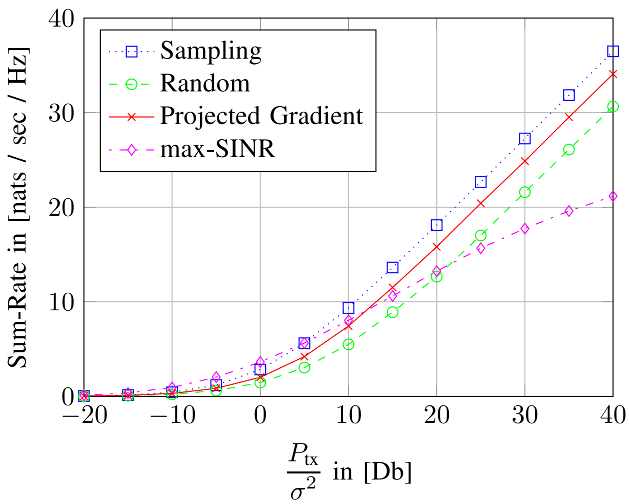

Fig. 2. Sum-Rate vs. SNR

E. Extensions

The suggested projected gradient approach for interference alignment can be easily extended to other differentiable utilities that are functions of the transmission rates. This allows for example to approximate an achievable rate region by using the weighted sum-rate as utility and varying the weights. Another example is to optimize for proportional fairness among the two users, where the utility is given by

$$\log\left(R_{11} + R_{12}\right) + \log\left(R_{21} + R_{22}\right).$$

Additionally the algorithm can be extended to the case where the number of antennas N at each transmitter and receiver is a multiple of three, by adjusting the dimensions accordingly.

IV. Simulation Results

We use 1000 realizations of complex Gaussian i.i.d channels and regard the average achieved sum-rates to evaluate the performance of the optimized interference alignment method, see Figure 2. The power budget of the two transmitters is the same $P_{\text{tx},1} = P_{\text{tx},2} = P_{\text{tx}}$. As comparison we include the max-SINR scheme presented in [3], which was actually developed for the interference channel where communication is pairwise. However, for the X channel it fails to find an aligned solution, as for multiple streams per transmitter the algorithm does not ensure that the filters at each transmitter are non-collinear. The work of [1] needs non-linear operations (dirty paper coding) at the transmitter and can therefore not be compared with solutions that build on linear filters. Finally we include the results for a random choice of $\boldsymbol{u}_{11}$ and $\boldsymbol{u}_{22}$, and for sampling $\boldsymbol{u}_{11}$ and $\boldsymbol{u}_{22}$ by using the best of 10000 random choices, which should be very close to the global optimum. In both cases the power allocation is done by the waterfilling rule. We can see that our approach has a significant gain compared to randomly choosing the aligned configuration. Finding the globally optimal solution is numerically to exhaustive, so our method that leading to local optimal solution is an attractive approach.

V. Conclusions

We were able to show a novel formulation for the sum-rate optimization in the MIMO X channel that guarantees an aligned solution, which is not found by algorithms building on alternating optimization. Applying a local optimization method, which converges within a few steps, results in significant gains compared to choosing a random alignment, while the global optimum can only be found by exhaustive search. For future research it would be interesting to see if similar parametrizations exist for other scenarios and if the projected gradient approach can be applied.

References

[1] M. Maddah-Ali, A. Motahari, and A. Khandani, “Communication over MIMO X channels: Interference alignment, decomposition, and performance analysis,” *IEEE Transactions on Information Theory*, vol. 54, no. 8, pp. 3457–3470, Aug. 2008.

[2] S. Jafar and S. Shamai, “Degrees of freedom region of the MIMO X channel,” *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 151–170, Jan. 2008.

[3] K. Gomadam, V. Cadambe, and S. Jafar, “Approaching the capacity of wireless networks through distributed interference alignment,” in *IEEE Global Telecommunications Conference, GLOBECOM 2008*, Dec. 2008.

[4] D. A. Schmidt, C. Shi, R. A. Berry, M. L. Honig, and W. Utschick, “Minimum mean squared error interference alignment,” in *Proceedings of the 43th Asilomar Conference on Signals, Systems and Computers*, Nov. 2009.

[5] C. Shi, D. Schmidt, R. Berry, M. Honig, and W. Utschick, “Distributed interference pricing for the MIMO interference channel,” in *IEEE International Conference on Communications, ICC 2009.*, June 2009.