Communication Over MIMO X Channels: Interference Alignment, Decomposition, and Performance Analysis

Mohammad Ali Maddah-Ali, *Member, IEEE*, Abolfazl Seyed Motahari, *Student Member, IEEE*, and Amir Keyvan Khandani, *Member, IEEE*

***Abstract*—In a multiple-antenna system with two transmitters and two receivers, a scenario of data communication, known as the X channel, is studied in which each receiver receives data from both transmitters. In this scenario, it is assumed that each transmitter is unaware of the other transmitter's data (noncooperative scenario). This system can be considered as a combination of two broadcast channels (from the transmitters' points of view) and two multiple-access channels (from the receivers' points of view). Taking advantage of both perspectives, two signaling schemes for such a scenario are developed. In these schemes, some linear filters are employed at the transmitters and at the receivers which decompose the system into either two noninterfering multiple-antenna broadcast subchannels or two noninterfering multiple-antenna multiple-access subchannels. The main objective in the design of the filters is to exploit the structure of the channel matrices to achieve the highest multiplexing gain (MG). It is shown that the proposed noncooperative signaling schemes outperform other known noncooperative schemes in terms of the achievable MG. In particular, it is shown that in some specific cases, the achieved MG is the same as the MG of the system if full cooperation is provided either between the transmitters or between the receivers.**

In the second part of the paper, it is shown that by using mixed design schemes, rather than decomposition schemes, and taking the statistical properties of the interference terms into account, the power offset of the system can be improved. The power offset represents the horizontal shift in the curve of the sum–rate versus the total power in decibels.

***Index Terms*—Degrees of freedom, interference alignment, interference channels, multiple-antenna systems, multiple-input multiple-output (MIMO) multiuser systems, MIMO X channels, multiplexing gain, noncooperative communication, power offset, space-division multiple access.**

Manuscript received June 14, 2007; revised November 20, 2007. This work was supported by Nortel and the corresponding matching funds by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Ontario Centres of Excellence (OCE). The material in this paper was presented in part at the IEEE International Symposium on Information Theory (ISIT), Seattle, WA, July 2006 and the 10th Canadian Workshop on Information Theory (CWIT) Edmonton, AB, Canada, June 2007 and reported in the Technical Report UW-ECE 2006-12, July 23, 2006, and the Technical Report UW-ECE 2006-27, December 2006, University of Waterloo, Waterloo, ON, Canada.

M. A. Maddah-Ali is with the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA 94720 USA (e-mail: maddah-a@eecs.berkeley.edu).

A. S. Motahari and A. K. Khandani are with the Coding and Signal Transmission Laboratory, Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: abolfazl@cst.uwaterloo.ca; khandani@cst.uwaterloo.ca).

Communicated by A. J. Grant, Associate Editor for Communications.

Color versions of Figures 6 and 7 in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIT.2008.926460

I. Introduction

WIRELESS technology has been advancing at an exponential rate, due to the increasing expectations for multiple-media services. This, in turn, necessitates the development of novel signaling techniques with high spectral efficiency. Using multiple antennas at both ends of wireless links is known as a unique solution to support high-data-rate communication [1], [2]. Multiple-antenna systems incorporate additional dimension of space to the transmission, resulting in a multiplicative increase in the overall throughput [2], [3]. The multiplicative increase in the rate is measured by a metric known as the multiplexing gain (MG), ρ, defined as the ratio of the sum–rate of the system, R, over the logarithm of the total power P_T in the high power regime, i.e.,

$$\rho = \lim_{P_T \to \infty} \frac{R}{\log_2(P_T)}. \quad (1)$$

It is widely known that in a point-to-point multiple-antenna system, with M transmit and N receive antennas, the MG is $\min(M, N)$[2]. In multiple-antenna multiple-user systems, when the full cooperation is provided at least at one side of the links (either among the transmitters or among the receivers), the system still enjoys a multiplicative increase in the throughput with the smaller value of the following two quantities: the total number of transmit antennas, and the total number of receive antennas. For example, in a multiple-access channel with two transmitters, with M_1 and M_2 antennas, and one receiver with N antennas, the MG is equal to $\min(M_1 + M_2, N)$[4]. Similarly, in a multiple-antenna broadcast channel, with one transmitter, equipped with M antennas, and two receivers, equipped with N_1 and N_2 antennas, the MG is equal to $\min(M, N_1 + N_2)$[4]. However, for the case that cooperation is not available, the performance of the system will be deteriorated due to the interference of the links over each other. For example, in a multiple-antenna interference channel with two transmitters and two receivers, each equipped with N antennas, the MG of the system is N[4].

Extensive research efforts have been devoted to the multiple-antenna interference channels. In [5], the capacity region of the multiple-input single-output (MISO) interference channel with strong interference (see [6]) and the capacity region of the single-input multiple-output (SIMO) interference channel with very strong interference (see [7]) are characterized. In [8], the

0018-9448/$25.00 © 2008 IEEE

superposition coding technique is utilized to derive an inner bound for the capacity of the multiple-input multiple-output (MIMO) interference channels. In [9], several numerical schemes are proposed to compute suboptimal transmit covariance matrices for the MIMO interference channels. In [4], the MG of the MIMO interference channel with general configuration for the number of transmit and receive antennas is derived. To increase the MG of such systems, the full cooperation among transmitters is proposed in [10], [11], which reduces the system to a single MIMO broadcast channel. To provide such a strong cooperation, an infinite capacity link connecting the transmitters, is presumed. In [12], the performance of single-antenna interference channels is evaluated, where the transmitters or receivers rely on the same channel, used for transmission, to provide cooperation. It is shown that the resulting MG is still one, i.e., this type of cooperation is not helpful in terms of the MG. In [4], a cooperation scheme in the shared communication channel for the MIMO interference systems is proposed and shown that such a scheme does not increase the MG.

In this paper, we propose a new signaling scenario in multiple-antenna systems with two transmitters and two receivers. In this scenario, each receiver receives data from both transmitters. It is assumed that neither the transmitters nor the receivers cooperate in signaling. In other words, each transmitter is unaware of the data of the other transmitter. Similarly, each receiver is unaware of the signal received by the other receiver. This scenario of signaling has several applications. For example, in a wireless system where two relay nodes are utilized to extend coverage area or in a system where two base stations provide different services to the users. This system can be considered as a combination of two broadcast channels (from the transmitters' points of view) and two multiple-access channels (from the receivers' points of view). By taking advantage of both perspectives, it is shown that by using some linear filters at the transmitters and the receivers, the system is decomposed to either two noninterfering multiple-antenna broadcast subchannels or two noninterfering multiple-antenna multiple-access subchannels. It is proven that such a scheme outperforms other known noncooperative schemes in terms of the achievable MG. In particular, it is shown that in the specific case that both receivers (transmitters) are equipped with N antennas, the total MG of $\rho = \lfloor \frac{4N}{3} \rfloor$ is achievable, where the two transmitters (receivers) have $\lceil \frac{\rho}{2} \rceil$ and $\lfloor \frac{\rho}{2} \rfloor$ antennas, respectively. Note that even if the full cooperation is provided either between the transmitters or between the receivers, the maximum MG is still ρ. Next, it is argued that such decomposition schemes result in some degradation (power offset) in the performance of the system. To overcome this problem, a design is proposed in which the signaling scheme is jointly designed for both subchannels (two broadcast or two multiple-access subchannels).

The authors proposed this scenario of signaling and established the possibility of achieving higher MG initially in [13]. Later in [14], we extended the scheme proposed in [13] to more general configurations for the number of transmit and receive antennas, and developed two signaling schemes based on: i) linear operations at the receivers and the dirty paper coding at the transmitters, and ii) linear operations at the transmitters and the successive decoding at the receivers. In addition, we introduced the concept of interference alignment for the first time as the main tool to achieve higher multiplexing gain (see [14, Sec. IV-B]). Then, in [15], the idea of interference alignment is elegantly used to show that zero-forcing scheme can achieve the multiplexing gain of the X channel. Furthermore, in [15], an upper bound on the MG of the X channels, where each transmitter and receiver is equipped with N antennas, is derived. In [16], the X channel with the partial and asymmetric cooperation among transmitters has been considered, and the MG of the system has been derived. Recently, in [17], it has been shown that by extending the X channel in time, the gap between the achievable MG and the upper bound proposed in [15] is closed.

The rest of the paper is organized as follows. In Section II, the system model is explained. In Section III, the two signaling schemes which decompose the system into two broadcast or two multiple-access subchannels are explained. The performance analysis of the scheme, including computing the MG and the power offset (for some special cases) is presented in Section IV. In Section V, the decomposition scheme is modified and a joint design for signaling scheme is proposed. Simulation results are presented in Section VI. Concluding remarks are presented in Section VII. *Notation:* All boldface letters indicate vectors (lower case) or matrices (upper case). $(\cdot)^\dagger$ denotes transpose conjugate operation, and $\mathcal{C}$ represents the set of complex numbers. $\mathcal{OC}^{M\times N}$ represents the set of all $M \times N$ complex matrices with mutually orthogonal and normal columns. $\boldsymbol{A} \perp \boldsymbol{B}$ means that each column of the matrix $\boldsymbol{A}$ is orthogonal to all columns of the matrix $\boldsymbol{B}$. The subspace spanned by columns of $\boldsymbol{A}$ is represented by $\Omega(\boldsymbol{A})$. The null space of the matrix $\boldsymbol{A}$ is denoted by $\mathrm{N}(\boldsymbol{A})$. Identity matrix is represented by $\boldsymbol{I}$. Adopted from MATLAB notation, $\boldsymbol{x}(i:j)$ denotes a vector including the entries i to j of the vector $\boldsymbol{x}$. The ith column of the matrix $\boldsymbol{A}$ is shown by $\boldsymbol{a}^{(i)}$.

II. Channel Model

We consider a MIMO system with two transmitters and two receivers. Transmitter t, $t = 1, 2$, is equipped with M_t antennas and receiver r, $r = 1, 2$, is equipped with N_r antennas. This configuration of antennas is shown by (M_1, M_2, N_1, N_2). For simplicity and without loss of generality, it is assumed that

$$M_1 \geq M_2 \text{ and } N_1 \geq N_2. \tag{2}$$

Assuming a flat-fading environment, the channel between transmitter t and receiver r is represented by the channel matrix $\boldsymbol{H}_{rt}$, where $\boldsymbol{H}_{rt} \in \mathcal{C}^{N_r \times M_t}$. The received vector $\boldsymbol{y}_r \in \mathcal{C}^{N_r \times 1}$ by receiver r, $r = 1, 2$, is given by

$$\begin{aligned} \boldsymbol{y}_1 &= \boldsymbol{H}_{11}\boldsymbol{s}_1 + \boldsymbol{H}_{12}\boldsymbol{s}_2 + \boldsymbol{w}_1 \\ \boldsymbol{y}_2 &= \boldsymbol{H}_{21}\boldsymbol{s}_1 + \boldsymbol{H}_{22}\boldsymbol{s}_2 + \boldsymbol{w}_2 \end{aligned} \tag{3}$$

where $\boldsymbol{s}_t \in \mathcal{C}^{M_t \times 1}$ represents the transmitted vector by transmitter t. The vector $\boldsymbol{w}_r \in \mathcal{C}^{N_r \times 1}$ is a white Gaussian noise with zero mean and identity covariance matrix. The power of $\boldsymbol{s}_t$ is subject to the constraint $\mathrm{Tr}(E[\boldsymbol{s}_t\boldsymbol{s}_t^\dagger]) \leq P_t$, $t = 1, 2$. P_T denotes the total transmit power, i.e., $P_T = P_1 + P_2$.

In the proposed scenario, the transmitter t sends μ_{1t} data streams to receiver 1 and μ_{2t} data streams to receiver 2.

Throughout the paper, we have the following assumptions.

- The perfect information of the all channel matrices $\boldsymbol{H}_{rt}$, $r,t = 1,2$ is available at both transmitters and at both receivers.
- Each transmitter is unaware of the data sent by the other transmitter, which means that there is no cooperation between transmitters. Similarly, receivers do not cooperate.

III. Decomposition Schemes

In what follows, we propose two signaling schemes depending on the values of (M_1, M_2, N_1, N_2). In the first scheme, by using linear transformations at the transmitters and at the receivers, the system is decomposed into two noninterfering broadcast subchannels. Therefore, we can use any signaling scheme, developed for the MIMO broadcast channels, over the resulting subchannels. As a dual of the first scheme, in the second scheme, linear transformations are utilized to decompose the system into two noninterfering multiple-access subchannels. It is shown that depending on the value of (M_1, M_2, N_1, N_2), one of the two schemes offer a higher MG.

In the rest of the paper, it is assumed that

$$M_1 < N_1 + N_2 \tag{4}$$

$$N_1 < M_1 + M_2. \tag{5}$$

Otherwise, if $M_1 \geq N_1 + N_2$, the maximum multiplexing gain of $N_1 + N_2$ is achievable by a simple broadcast channel formed by the first transmitter and the two receivers. Similarly, if $N_1 \geq M_1 + M_2$, then the maximum multiplexing gain of $M_1 + M_2$ is achievable by a simple multiple-access channel including the two transmitters and the first receiver. The two signaling schemes presented in this paper cover all the possibilities for the number of transmit and receive antennas, excluding the aforementioned cases. The optimality is proven for some special cases of practical interest.

To attain the highest MG, we take advantage of the null spaces of the direct or cross links.

Defintion 1: We call a system *irreducible*, if

$$\text{Irreducible Type I:} \quad N_1 \geq N_2 \geq M_1 \geq M_2 \tag{6}$$

or

$$\text{Irreducible Type II:} \quad M_1 \geq M_2 \geq N_1 \geq N_2. \tag{7}$$

Otherwise the system is called *reducible*.

Unlike the irreducible systems, a portion of the achieved MG in a reducible X channel is attained through exploiting the null spaces of the direct or cross links. In the reducible systems, the null spaces of the links provide the possibility to increase the number of data streams sent from one of the transmitters to one of the receivers, without imposing any interference on the other receiver or restricting the signaling space of the other transmitter. By excluding null spaces utilized to increase the MG, the system is reduced to an equivalent system with (M_1', M_2', N_1', N_2') antennas, where $(M_1', M_2', N_1', N_2') \leq (M_1, M_2, N_1, N_2)$. As will be explained later, the null spaces of the links in the reducible systems are exploited to the extent that the equivalent (reduced) system is not reducible anymore.

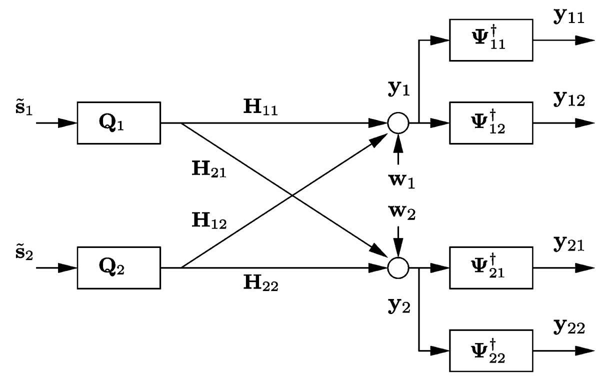

Fig. 1. Scheme one: decomposition of the system into two broadcast subchannels.

Defintion 2: If the reduced X channel satisfies the condition of the Type I irreducible systems, i.e., $N_1' \geq N_2' \geq M_1' \geq M_2'$, the original system is called reducible to Type I. Similarly, if $M_1' \geq M_2' \geq N_1' \geq N_2'$, the original system is called reducible to Type II.

In what follows, it is shown that the Type I irreducible systems and the reducible systems to Type I can be decomposed into two noninterfering broadcast sub-channels. Moreover, it is shown that the Type II irreducible systems, and the reducible systems to Type II can be decomposed into two noninterfering multiple-access subchannels.

We define μ_{rt}', $r,t = 1,2$, as the number of data streams transmitted from transmitter t to receiver r, *excluding* the number of extra data streams attained through exploiting the null spaces of the links. In other words, μ_{rt}' represents the number of data streams in the equivalent (reduced) channel.

A. Scheme I—Decomposition of the System Into Two Broadcast Subchannels

As depicted in Fig. 1, in this scheme, the transmit filter $\boldsymbol{Q}_t \in \mathcal{OC}^{M_t \times (\mu_{1t}+\mu_{2t})}$ is employed at transmitter t, $t = 1,2$. Therefore, the transmitted vectors $\boldsymbol{s}_t$, $t = 1,2$, are equal to

$$\boldsymbol{s}_t = \boldsymbol{Q}_t \tilde{\boldsymbol{s}}_t \tag{8}$$

where $\tilde{\boldsymbol{s}}_t \in \mathcal{C}^{(\mu_{1t}+\mu_{2t})\times 1}$ contains μ_{1t} data streams for receiver one and μ_{2t} data streams for receiver two. The transmit filters $\boldsymbol{Q}_t$, $t = 1,2$, have two functionalities: i) Confining the transmit signal from transmitter t to a $(\mu_{1t}+\mu_{2t})$-dimensional subspace which provides the possibility to decompose the system into two broadcast subchannels by using linear filters at the receivers. ii) Exploiting the null spaces of the channel matrices to achieve the highest multiplexing gain.

At each receiver, two parallel receive filters are employed. The received vector $\boldsymbol{y}_1$ is passed through the filter $\mathbf{\Psi}_{11}^\dagger$, which is used to null out the signal coming from the second transmitter. The μ_{11} data streams, sent by transmitter one intended to receiver one, can be decoded from $\boldsymbol{y}_{11}$, the output of $\mathbf{\Psi}_{11}^\dagger$. Similarly, to decode μ_{12} data streams, sent by transmitter two

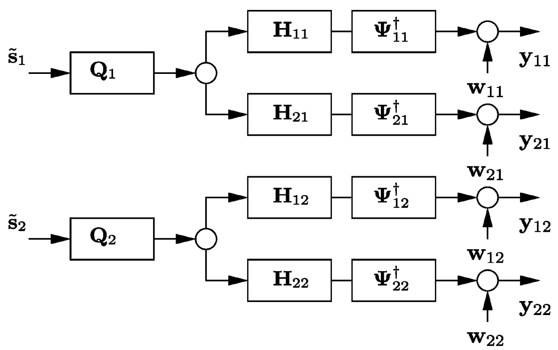

Fig. 2. Scheme one: The resulting noninterfering MIMO broadcast subchannels.

to receiver one, the received vector $\boldsymbol{y}_1$ is passed through the receive filter $\boldsymbol{\Psi}_{12}^{\dagger}$, which is used to null out the signal coming from transmitter one. Receiver two has a similar structure with parallel receive filters $\boldsymbol{\Psi}_{21}^{\dagger}$ and $\boldsymbol{\Psi}_{22}^{\dagger}$. Later, it is shown that if μ_{rt}, $r,t=1,2$, satisfy a set of inequalities, then it is possible to design $\boldsymbol{Q}_t$ and $\boldsymbol{\Psi}_{rt}$ to meet the desired features explained earlier. It means that the system is decomposed into two noninterfering MIMO broadcast sub-channels (see Fig. 2).

Next, we explain how to select the design parameters including the number of data streams μ_{rt}, $r,t=1,2$ and the transmit/receive filters. The primary objective is to prevent the saturation of the rate of each stream in the high signal-to-noise ratio (SNR) regime. In other words, the MG of the system is $\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22}$.

The integer variables ζ_{rt}, $r,t=1,2$, defined as follows, will be useful in our subsequent discussions:

- ζ_{11} denotes the dimension of $\Omega(\boldsymbol{H}_{12}\boldsymbol{Q}_2)$;
- ζ_{21} denotes the dimension of $\Omega(\boldsymbol{H}_{22}\boldsymbol{Q}_2)$;
- ζ_{12} denotes the dimension of $\Omega(\boldsymbol{H}_{11}\boldsymbol{Q}_1)$;
- ζ_{22} denotes the dimension of $\Omega(\boldsymbol{H}_{21}\boldsymbol{Q}_1)$.

In the sequel, we categorize the design scheme into the four general cases depending on (M_1,M_2,N_1,N_2), where in all cases, the system is either irreducible Type I or reducible to Type I. To facilitate the derivations, we use the auxiliary variables M'_t, N'_r, and μ'_{rt}, for $r,t=1,2$. As will be explain later, for each case, M'_t and N'_r are computed directly as a function of M_t and N_r for $r,t=1,2$. Then, μ'_{rt}, $r,t=1,2$, are selected such that the following constraints are satisfied:

$$\mu'_{11}:\quad \mu'_{11}+\mu'_{12}+\mu'_{22}\le N'_1 \tag{9}$$
$$\mu'_{12}:\quad \mu'_{12}+\mu'_{11}+\mu'_{21}\le N'_1 \tag{10}$$
$$\mu'_{22}:\quad \mu'_{22}+\mu'_{21}+\mu'_{11}\le N'_2 \tag{11}$$
$$\mu'_{21}:\quad \mu'_{21}+\mu'_{22}+\mu'_{12}\le N'_2 \tag{12}$$
$$\mu'_{11}+\mu'_{21}\le M'_1 \tag{13}$$
$$\mu'_{22}+\mu'_{12}\le M'_2. \tag{14}$$

Each of the first four inequalities corresponds to one of the parameters μ'_{rt}, $r,t=1,2$, in the sense that if μ'_{rt}, $r,t=1,2$, is zero, the corresponding inequality is removed from the set of constraints. After choosing μ'_{rt}, $r,t=1,2$, for each case, μ_{rt}, $r,t=1,2$, are computed as function of μ'_{rt}, $r,t=1,2$, as will be explained later. Note that we have many options to choose μ'_{rt}, $r,t=1,2$. It is shown that as long as the integers μ'_{rt}, $r,t=1,2$, satisfy (9)–(14), the system achieves the MG of $\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22}$. However, it turns out that to achieve the highest MG, μ'_{rt}, $r,t=1,2$, should be selected such that $\mu'_{11}+\mu'_{12}+\mu'_{21}+\mu'_{22}$ is maximum.

Next, for each of the four cases, we explain the following.

1) How to compute the auxiliary variables M'_t and N'_r as a function of M_t and N_r, $r,t=1,2$.
2) After choosing μ'_{rt}, $r,t=1,2$, satisfying (9)–(14), how to compute μ_{rt}, $r,t=1,2$.
3) How to choose the transmit filters $\boldsymbol{Q}_t$, $t=1,2$.
4) How to compute ζ_{rt}, $r,t=1,2$.

Having completed these steps, the procedure of computing the receive filters $\boldsymbol{\Psi}_{rt}^{\dagger}$, $r,t=1,2$, is similar for all cases. Later, we will show that this scheme decomposes the system into two noninterfering broadcast subchannels.

Scheme I–Case I: $N_1\ge N_2\ge M_1\ge M_2$: In this case, the system is irreducible. Therefore, the equivalent system is the same as the original system i.e., $N'_r=N_r$, $r=1,2$, and $M'_t=M_t$, $t=1,2$.

Using the above parameters, we choose μ'_{rt}, $r,t=1,2$, subject to (9)–(14) constraints. Since we do not exploit the null space of any of the links to transmit data streams, μ_{rt} is the same as μ'_{rt}, i.e., $\mu_{rt}=\mu'_{rt}$, $r,t=1,2$. In this case, $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ are randomly chosen from $\mathcal{OC}^{M_1\times(\mu_{11}+\mu_{21})}$ and $\mathcal{OC}^{M_2\times(\mu_{12}+\mu_{22})}$, respectively.

Regarding the definition of ζ_{rt}, $r,t=1,2$, it is easy to see that

$$\begin{aligned}\zeta_{11}&=\mu_{12}+\mu_{22}, &\zeta_{12}&=\mu_{11}+\mu_{21}\\ \zeta_{21}&=\mu_{12}+\mu_{22}, &\zeta_{22}&=\mu_{11}+\mu_{21}.\end{aligned} \tag{15}$$

Scheme I–Case II: $N_1\ge M_1>N_2\ge M_2$: In this case, at transmitter one, (M_1-N_2)-dimensional subspace $\mathrm{N}(\boldsymbol{H}_{21})$ is exploited to transmit M_1-N_2 data streams from transmitter one to receiver one without imposing any interference at receiver two. In other words, while the component of $\boldsymbol{s}_1$ in $\mathrm{N}(\boldsymbol{H}_{21})$ does not impose any interference at receiver two, it provides the possibility to increase the number of data streams sent from transmitter one to receiver one by M_1-N_2. Let us exclude the (M_1-N_2)-dimensional subspace $\mathrm{N}(\boldsymbol{H}_{21})$ from the available space at transmitter one. In addition, let us exclude the (M_1-N_2)-dimensional subspace $\Omega(\boldsymbol{H}_{11}\mathrm{N}(\boldsymbol{H}_{21}))$ from the available space at receiver one. Then, the system is reduced to an X channel with equivalent antennas

$$(M'_1,M'_2,N'_1,N'_2)=\Big(\{M_1-\{M_1-N_2\},M_2,N_1-\{M_1-N_2\},N_2\Big)$$

or

$$N'_1=N_1+N_2-M_1,\ N'_2=N_2,\ M'_1=N_2,\ M'_2=M_2. \tag{16}$$

Clearly, $N'_1\ge N'_2\ge M'_1\ge M'_2$, therefore the original system is reducible to Type I.

Let us select μ'_{rt}, $r,t=1,2$, subject to (9)–(14) constraints. μ'_{rt}, $r,t=1,2$, give us the number of data streams in the reduced X channel, excluding the M_1-N_2 data streams, sent from transmitter one to receiver one relying on $\mathrm{N}(\boldsymbol{H}_{21})$. Clearly, the numbers of data streams in the original system are computed as,

$$\begin{aligned}\mu_{11} &= \mu'_{11} + M_1 - N_2, \quad \mu_{12} = \mu'_{12},\\ \mu_{21} &= \mu'_{21}, \quad \mu_{22} = \mu'_{22}.\end{aligned} \tag{17}$$

$\boldsymbol{Q}_1$ is chosen as

$$\boldsymbol{Q}_1 \in \mathcal{OC}^{M_1\times(\mu_{11}+\mu_{21})}, \quad \boldsymbol{Q}_1 = [\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2] \tag{18}$$

where

$$\boldsymbol{\Sigma}_1 \in \mathcal{OC}^{M_1\times(N_1-M_2)}, \quad \boldsymbol{\Sigma}_1 \in \mathrm{N}(\boldsymbol{H}_{21}) \tag{19}$$
$$\boldsymbol{\Sigma}_2 \in \mathcal{OC}^{M_1\times(\mu'_{11}+\mu_{21})}, \quad \boldsymbol{\Sigma}_2 \perp \boldsymbol{\Sigma}_1. \tag{20}$$

Such a structure for $\boldsymbol{\Sigma}_1$ guarantees the full usage of $\mathrm{N}(\boldsymbol{H}_{21})$ for signaling.

$\boldsymbol{Q}_2$ is randomly chosen from $\mathcal{OC}^{M_2\times(\mu_{12}+\mu_{22})}$.

It is easy to see that

$$\zeta_{11} = \mu_{12}+\mu_{22},\ \zeta_{12} = \mu_{11}+\mu_{21},\ \zeta_{21} = \mu_{12}+\mu_{22},\ \zeta_{22} = \mu'_{11}+\mu_{21}. \tag{21}$$

Scheme I–Case III: $\underline{N_1 \geq M_1 \geq M_2 > N_2}$ *and* $\underline{N_1 + N_2 \geq M_1 + M_2}$*:*

In this case

i) at transmitter one, $(M_1 - N_2)$-dimensional subspace $\mathrm{N}(\boldsymbol{H}_{21})$ is utilized to increase the number data streams sent from transmitter one to receiver one by $M_1 - N_2$ without imposing interference at receiver two;

ii) at transmitter two, $(M_2 - N_2)$-dimensional subspace $\mathrm{N}(\boldsymbol{H}_{22})$ is utilized to increase the number data streams sent from transmitter two to receiver one by $M_2 - N_2$ without imposing interference at receiver two.

We exclude

i) $(M_1 - N_2)$-dimensional subspace $\mathrm{N}(\boldsymbol{H}_{21})$ from the signaling space at transmitter one;

ii) $(M_2 - N_2)$-dimensional subspace $\mathrm{N}(\boldsymbol{H}_{22})$ from the signaling space at transmitter two;

iii) $(M_2 - N_1) + (M_2 - N_2)$-dimensional subspace $\Omega(\boldsymbol{H}_{11}\mathrm{N}(\boldsymbol{H}_{21})) \cup \Omega(\boldsymbol{H}_{12}\mathrm{N}(\boldsymbol{H}_{22}))$ from the signaling space at receiver one.

Then, the reduced system is an equivalent X channel with (M'_1, M'_2, N'_1, N'_2), where

$$N'_1 = N_1 + 2N_2 - M_1 - M_2,\ N'_2 = N_2,\ M'_1 = N_2,\ M'_2 = N_2 \tag{22}$$

where $N'_1 \geq N'_2 \geq M'_1 \geq M'_2$. Therefore, the original system is reducible to Type I. The number of data streams in the equivalent channel, μ'_{rt}, $r, t = 1, 2$, are selected subject to (9)–(14) constraints. Then, we have

$$\mu_{11} = \mu'_{11} + M_1 - N_2,\ \mu_{12} = \mu'_{12} + M_2 - N_2,\ \mu_{21} = \mu'_{21},\ \mu_{22} = \mu'_{22}. \tag{23}$$

$\boldsymbol{Q}_1$ is chosen as

$$\boldsymbol{Q}_1 \in \mathcal{OC}^{M_1\times(\mu_{11}+\mu_{21})}, \quad \boldsymbol{Q}_1 = [\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2] \tag{24}$$

where

$$\boldsymbol{\Sigma}_1 \in \mathcal{OC}^{M_1\times(M_1-N_2)}, \quad \boldsymbol{\Sigma}_1 \in \mathrm{N}(\boldsymbol{H}_{21}) \tag{25}$$
$$\boldsymbol{\Sigma}_2 \in \mathcal{OC}^{M_1\times(\mu'_{11}+\mu_{21})}, \quad \boldsymbol{\Sigma}_2 \perp \boldsymbol{\Sigma}_1. \tag{26}$$

$\boldsymbol{Q}_2$ is chosen as

$$\boldsymbol{Q}_2 \in \mathcal{OC}^{M_2\times(\mu_{12}+\mu_{22})}, \quad \boldsymbol{Q}_2 = [\boldsymbol{\Sigma}_3, \boldsymbol{\Sigma}_4] \tag{27}$$

where

$$\boldsymbol{\Sigma}_3 \in \mathcal{OC}^{M_2\times(M_2-N_2)}, \quad \boldsymbol{\Sigma}_3 \in \mathrm{N}(\boldsymbol{H}_{22}) \tag{28}$$
$$\boldsymbol{\Sigma}_4 \in \mathcal{OC}^{M_2\times(\mu'_{12}+\mu_{22})}, \quad \boldsymbol{\Sigma}_4 \perp \boldsymbol{\Sigma}_3. \tag{29}$$

It is easy to see that

$$\zeta_{11} = \mu_{12}+\mu_{22},\ \zeta_{12} = \mu_{11}+\mu_{21},\ \zeta_{21} = \mu'_{12}+\mu_{22},\ \zeta_{22} = \mu'_{11}+\mu_{21}. \tag{30}$$

Scheme I–Case IV: $\underline{M_1 > N_1 \geq N_2 \geq M_2}$ *and* $\underline{N_1 + N_2 \geq M_1 + M_2}$*:*

In this case, at transmitter one, i) $(M_1 - N_2)$-dimensional subspace $\mathrm{N}(\boldsymbol{H}_{21})$ is utilized to increase the number data streams sent from transmitter one to receiver one by $M_1 - N_2$ without imposing interference at receiver two, ii) $(M_1 - N_1)$-dimensional subspace $\mathrm{N}(\boldsymbol{H}_{11})$ is exploited to increase the number data streams from transmitter one to receiver two by $M_1 - N_1$, without imposing interference at receiver two. By excluding the utilized subspaces at transmitter one, receiver one, and receiver two, the equivalent system is an X channel with (M'_1, M'_2, N'_1, N'_2) where

$$N'_1 = N_1+N_2-M_1,\ N'_2 = N_1+N_2-M_1,\ M'_1 = N_1+N_2-M_1,\ M'_2 = M_2. \tag{31}$$

It is easy to see that $N'_1 \geq N'_2 \geq M'_1 \geq M'_2$. Therefore, the original system is reducible to Type I. μ'_{rt}, $r, t = 1, 2$, are selected subject to (9)–(14) constraints. Then

$$\mu_{11} = \mu'_{11} + M_1 - N_2,\ \mu_{12} = \mu'_{12},\ \mu_{21} = \mu'_{21} + M_1 - N_1,\ \mu_{22} = \mu'_{22}. \tag{32}$$

In addition, $\boldsymbol{Q}_1$ is chosen as

$$\boldsymbol{Q}_1 \in \mathcal{OC}^{M_1\times(\mu_{11}+\mu_{21})}, \quad \boldsymbol{Q}_1 = [\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2] \tag{33}$$

where

$$\boldsymbol{\Sigma}_1 \in \mathcal{OC}^{M_1\times(M_1-N_2+M_1-N_2)}, \quad \boldsymbol{\Sigma}_1 \in \mathrm{N}(\boldsymbol{H}_{21}) \cup \mathrm{N}(\boldsymbol{H}_{11}) \tag{34}$$
$$\boldsymbol{\Sigma}_2 \in \mathcal{OC}^{M_1\times(\mu'_{11}+\mu'_{21})}, \quad \boldsymbol{\Sigma}_2 \perp \boldsymbol{\Sigma}_1. \tag{35}$$

$\boldsymbol{Q}_2$ is randomly chosen from $\mathcal{OC}^{M_2\times(\mu_{12}+\mu_{22})}$.

It is easy to see that

$$\zeta_{11}=\mu_{12}+\mu_{22},\ \zeta_{12}=\mu_{11}+\mu_{21}',\ \zeta_{21}=\mu_{12}+\mu_{22},\ \zeta_{22}=\mu_{11}'+\mu_{21}. \quad (36)$$

The next steps of the algorithm are the same for all of the aforementioned cases. We define

$$\widetilde{\boldsymbol{H}}_{rt}=\boldsymbol{H}_{rt}\boldsymbol{Q}_t, \quad r,t=1,2. \quad (37)$$

$\boldsymbol{\Psi}_{rt}\in\mathcal{OC}^{N_t\times(N_t-\zeta_{rt})}$, $r,t=1,2$, are chosen such that

$$\boldsymbol{\Psi}_{11}\perp\widetilde{\boldsymbol{H}}_{12} \quad (38)$$
$$\boldsymbol{\Psi}_{12}\perp\widetilde{\boldsymbol{H}}_{11} \quad (39)$$
$$\boldsymbol{\Psi}_{21}\perp\widetilde{\boldsymbol{H}}_{22} \quad (40)$$
$$\boldsymbol{\Psi}_{22}\perp\widetilde{\boldsymbol{H}}_{21}. \quad (41)$$

According to the definition of ζ_{rt}, one can always choose such matrices. Clearly, any signal sent by transmitter one does not pass through the filters $\boldsymbol{\Psi}_{12}^{\dagger}$ and $\boldsymbol{\Psi}_{22}^{\dagger}$. Similarly, any signal sent by transmitter two does not pass through the filters $\boldsymbol{\Psi}_{21}^{\dagger}$ and $\boldsymbol{\Psi}_{11}^{\dagger}$.

We define

$$\overline{\boldsymbol{H}}_{rt}=\boldsymbol{\Psi}_{rt}^{\dagger}\widetilde{\boldsymbol{H}}_{rt}, \quad r,t=1,2 \quad (42)$$
$$\boldsymbol{w}_{rt}=\boldsymbol{\Psi}_{rt}^{\dagger}\boldsymbol{w}_r, \quad r,t=1,2 \quad (43)$$

and

$$\boldsymbol{y}_{rt}=\boldsymbol{\Psi}_{rt}^{\dagger}\boldsymbol{y}_r, \quad r,t=1,2. \quad (44)$$

Therefore, the system is decomposed into two noninterfering broadcast channels. The MIMO broadcast channel viewed from transmitter 1 is modeled by (see Fig. 2)

$$\begin{cases}\boldsymbol{y}_{11}=\overline{\boldsymbol{H}}_{11}\tilde{\boldsymbol{s}}_1+\boldsymbol{w}_{11}\\ \boldsymbol{y}_{21}=\overline{\boldsymbol{H}}_{21}\tilde{\boldsymbol{s}}_1+\boldsymbol{w}_{21}\end{cases} \quad (45)$$

and the MIMO broadcast channel viewed from transmitter two is modeled by (see Fig. 2)

$$\begin{cases}\boldsymbol{y}_{12}=\overline{\boldsymbol{H}}_{12}\tilde{\boldsymbol{s}}_2+\boldsymbol{w}_{12}\\ \boldsymbol{y}_{22}=\overline{\boldsymbol{H}}_{22}\tilde{\boldsymbol{s}}_2+\boldsymbol{w}_{22}.\end{cases} \quad (46)$$

B. *Scheme II—Decomposition of the System Into Two Multiple-Access Subchannels*

This scheme is indeed the dual of the scheme one, detailed in Section III-A (see Figs. 3 and 4). The discussion related to Scheme II follows a path similar to what is discussed above for Scheme I as given in Appendix I.

IV. Performance Evaluation

The decomposition schemes, presented in Section III, simplify the procedure of the performance evaluation for the X channels in the high-SNR regime. In what follows, the MG of the X channel is studied. In addition, for some special cases, a metric known as *power offset* is evaluated.

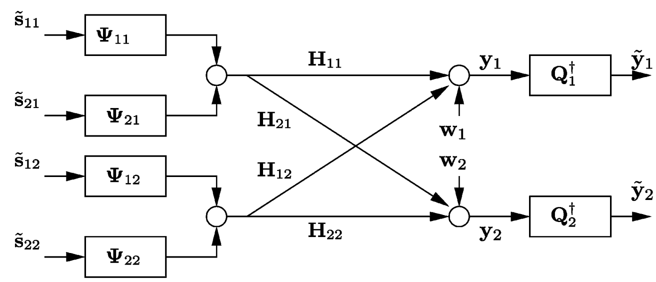

Fig. 3. Scheme II: Decomposition of the system into two multiple-access subchannels.

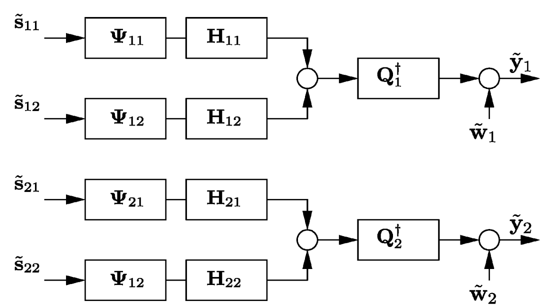

Fig. 4. Scheme II: The resulting noninterfering MIMO multiple-access subchannels.

A. *Multiplexing Gain*

Theorem 3: The MIMO X channel with (M_1,M_2,N_1,N_2) antennas, decomposed into two noninterfering broadcast or multiple-access subchannels, achieves the multiplexing gain of $\mu_{11}+\mu_{21}+\mu_{12}+\mu_{22}$, if $\mu_{rt}, r,t=1,2$, are selected according to the schemes presented in Section III.

Proof: As explained in Section III-A, the X channel is decomposed into two noninterfering broadcast subchannels(45) and (46). The first broadcast subchannel is formed with the channel matrices $\overline{\boldsymbol{H}}_{11}\in\mathcal{C}^{(\mu_{11}+\mu_{21})\times(N_1-\zeta_{11})}$ and $\overline{\boldsymbol{H}}_{21}\in\mathcal{C}^{(\mu_{11}+\mu_{21})\times(N_2-\zeta_{21})}$. The inequalities (9) and (12) guarantee that $N_1-\zeta_{11}\geq\mu_{11}$ and $N_2-\zeta_{21}\geq\mu_{21}$. Note that the input of the MIMO broadcast subchannel viewed from transmitter one is $\tilde{\boldsymbol{s}}_1$. On the other hand, $\boldsymbol{Q}_1$ is chosen such that $\boldsymbol{Q}_1^{\dagger}\boldsymbol{Q}_1=\boldsymbol{I}$. Therefore, $E(\tilde{\boldsymbol{s}}_1^{\dagger}\tilde{\boldsymbol{s}}_1)=E(\boldsymbol{s}_1^{\dagger}\boldsymbol{s}_1)$. Hence, the power constraints on the input signals is the same as the power constraint on $\tilde{\boldsymbol{s}}_1$. Consequently, we are free to choose any covariance matrix for the input vector $\tilde{\boldsymbol{s}}_1$, subject to $\mathrm{Tr}\, E(\tilde{\boldsymbol{s}}_1\tilde{\boldsymbol{s}}_1^{\dagger})\leq P_1$. Therefore, as long as the matrix $[\overline{\boldsymbol{H}}_{11}^{\dagger},\overline{\boldsymbol{H}}_{21}^{\dagger}]$ is full rank, the broadcast subchannel achieves the MG of $\mu_{11}+\mu_{21}$ by sending μ_{11} data streams to receiver one and μ_{21} data streams to receiver two. It is easy to see that the $[\overline{\boldsymbol{H}}_{11}^{\dagger},\overline{\boldsymbol{H}}_{21}^{\dagger}]$ is full rank with probability one. Similarly, the second broadcast subchannel is formed with the channel matrices $\overline{\boldsymbol{H}}_{12}\in\mathcal{C}^{(\mu_{12}+\mu_{22})\times(N_1-\zeta_{12})}$, and $\overline{\boldsymbol{H}}_{22}\in\mathcal{C}^{(\mu_{12}+\mu_{22})\times(N_2-\zeta_{22})}$. Constraints (11) and (12), respectively, guarantee that $N_1-\zeta_{21}\geq\mu_{21}$ and $N_2-\zeta_{22}\geq\mu_{22}$. Therefore, as long as the matrix $[\overline{\boldsymbol{H}}_{12}^{\dagger},\overline{\boldsymbol{H}}_{22}^{\dagger}]$ is full rank, the second broadcast subchannel achieves the MG of $\mu_{12}+\mu_{22}$ by sending μ_{12} data streams to receiver one and μ_{22} data streams to receiver two.

A similar arguments are valid for the scheme presented in Section III-B. □

Next, the MG of some special cases is computed in a closed form.

Corollary 4: For the special case of $N_1 = N_2 = N$ in the scheme of Section III-A, the MG of $\rho = \left\lfloor \frac{4N}{3} \right\rfloor$ is achievable where the total number of transmit antennas is equal to ρ, which are divided between transmitters as $M_1 = \left\lceil \frac{\rho}{2} \right\rceil$ and $M_2 = \left\lfloor \frac{\rho}{2} \right\rfloor$.

Proof: By direct verification in the Scheme I–Case I. □

Corollary 5: In the special case of $M_1 = M_2 = M$ in the scheme presented in Section III-B, the MG of $\rho = \left\lfloor \frac{4M}{3} \right\rfloor$ is achievable where the total number of receive antennas is equal to ρ, which are divided between receivers as $N_1 = \left\lceil \frac{\rho}{2} \right\rceil$ and $N_2 = \left\lfloor \frac{\rho}{2} \right\rfloor$.

Proof: By direct verification in the Scheme II–Case I. □

Regarding Theorem 3, the MG of the X channel outperforms the MG of the interference channel with the same number of antennas. For example, the MGs of a X channels with $(3,3,3,3)$, $(4,3,4,3)$, $(9,5,8,7)$ antennas are $4, 5$, and 11, respectively, while the MGs of the interference channels with the same number of antennas are, respectively, $3, 4$, and 9 [4]. For all the cases listed in Corollaries 4 and 5, the MG of the X channel is the same as the MG of the system with full cooperation between transmitters or between receivers. For example, the multiplexing gains of the X channels with $(2,2,3,3)$, $(3,3,2,2)$, $(3,2,4,4)$, and $(3,3,5,5)$ antennas are, respectively, $4, 4, 5$, and 6.

The improvement in MG of the X channels as compared to the interference channels can be attributed to two phenomena as explained next.

1) *Interference Alignment:* For simplicity, we consider an X channel with $(2,2,3,3)$ antennas, and assume that transmitter t sends one data stream d_{rt} to receiver r, $r = 1, 2$. Therefore, there are four data streams in the shared wireless medium. At receiver one, we are interested in decoding d_{11} and d_{12}, while d_{22} and d_{21} are treated as interference. The signaling scheme is designed such that at the receiver one terminal, the interference terms d_{21} and d_{22} are received in the directions for which the distractive components are along each other. Therefore, at receiver one with three antennas, one direction is occupied with the destructive component of both interference terms d_{21} and d_{22}, while we have two interference-free dimensions to receive d_{11} and d_{12}. The design scheme provides similar condition at the receiver two terminal, while d_{22} and d_{21} are desired data streams and d_{22} and d_{21} are interference terms. Such overlaps of interference terms in each receiver save the available spatial dimensions to exploit the highest MG.

2) *Maximizing the Possibility of Cooperation Among Data Streams:* It is well known that the MG for a point-to-pint MIMO channel, a MIMO broadcast channel, and a MIMO multiple-access channel is the same, as long as in all three systems we have the same total number of transmit antennas and the same total number of receive antennas. The immediate conclusion is that to attain the maximum MG, the cooperation at one side of the communication link is enough.

Now, consider an interference channel with $M_1 = M_2 = 2$ and $N_1 = N_2 = 3$, and assume that two data streams d_1 and d_2 are sent from transmitter one to receiver one and two data streams d_3 and d_4 are sent from transmitter two to receiver two. In this scenario, the data streams d_1 and d_2 have the possibility to cooperate at two points: i) at transmitter one, and ii) at receiver one. Similarly, the data streams d_3 and d_4 have the possibility to cooperate at two points: i) at transmitter two, and ii) at receiver two. Regarding the aforementioned discussion, the system does not gain MG from the provided cooperation for d_1 and d_2 at both transmitter one and receiver one. Similar argument is valid for d_3 and d_4. However, the performance of the system is deteriorated because there is no possibility to cooperate between (d_1, d_2) and (d_3, d_4).

Let us consider an X channels with $(2,2,3,3)$ antennas. In the X channels, the cooperation between d_{11} and d_{21} is provided at transmitter one, and the cooperation between d_{12} and d_{22} is provided at transmitter two. Similarly, the cooperation between d_{11} and d_{12} is provided at receiver one, and the cooperation between d_{21} and d_{22} is provided at receiver two.

B. Power Offset

In Corollaries 4 and 5, some special cases are listed for which the MG of the X channel is the same as the MG of a point-to-point MIMO system resulting from full cooperation between transmitters and/or between receivers. However, it does not mean that the system does not gain any improvement through cooperation. The gain of the cooperation is reflected in a metric known as the *power offset*. The power offset is defined as the negative of the zero-order term in the expansion of the sum–rate with respect to the total power, normalized with multiplexing gain, i.e.,

$$R = \rho(\log_2(P_T) - \mathcal{L}_\infty) + o(1) \tag{47}$$

where P_T denotes the total power, and $\mathcal{L}_\infty$ denotes the power offset in 3-dB units (see Fig. 5). In this definition, it is assumed that the noise is normalized as in system model (3). The power offset was first introduced in [18] to evaluate the performance of the different code-division multiple-access (CDMA) schemes. Later, the power offset for MIMO channels in [19] and some special cases of MIMO broadcast channels in [20] were computed. In what follows, the result of [20] is adopted to compute the power offset of some special cases of MIMO X channels.

Theorem 6: In an X channel with $(M_1, M_2, N_1, N_2) = (2\bar{k}, 2\bar{k}, 3\bar{k}, 3\bar{k})$ antennas ($\bar{k}$ is a positive integer number), where the entries of channel matrices have Rayleigh distribution, if the decomposition scheme is employed, the power offset is equal to

$$\mathcal{L}_\infty(M_1, M_2, N_1, N_2) = \mathcal{L}_\infty(2\bar{k}, 2\bar{k}) - \frac{1}{2}(\log_2(\alpha) + \log_2(1-\alpha)) \tag{48}$$

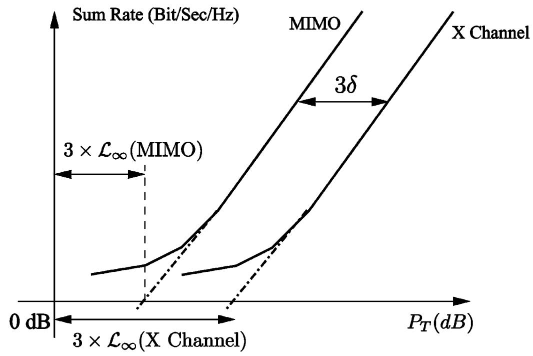

Fig. 5. Power offset of the MIMO point-to-point channels and the MIMO X channel.

in 3-dB units, where $P_1 = \alpha P_T, P_2 = (1-\alpha)P_T, 0 \le \alpha \le 1$

$$\mathcal{L}_\infty(M,N) = \log_2 M + \frac{1}{\ln(2)}\left(\bar{\gamma} + 1 - \sum_{i=1}^{\tilde{M}-\tilde{N}} \frac{1}{i} - \frac{\tilde{M}}{\tilde{N}} \sum_{i=\tilde{M}-\tilde{N}+1}^{\tilde{M}} \frac{1}{i}\right) \tag{49}$$

$\bar{\gamma} = 0.5772$, $\tilde{M} = \max\{M,N\}$, and $\tilde{N} = \min\{M,N\}$. Furthermore, the power offset of the X channel with $(2\bar{k}, 2\bar{k}, 3\bar{k}, 3\bar{k})$ antennas with respect to a MIMO Rayleigh channel with $4\bar{k}$ transmit antennas and $6\bar{k}$ receive antennas is equal to

$$\delta = \frac{3}{2\ln(2)} \sum_{i=2\bar{k}+1}^{6\bar{k}} \frac{1}{i} - 1 - \frac{1}{2}\left(\log_2(\alpha) + \log_2(1-\alpha)\right) \tag{50}$$

in 3-dB units.

Proof: In this case, the transmit filter $\boldsymbol{Q}_1$, is randomly chosen from $\mathcal{OC}^{2\bar{k}\times 2\bar{k}}$, independent of $\boldsymbol{H}_{11}$ and $\boldsymbol{H}_{21}$. In addition, the receive filters $\boldsymbol{\Psi}_{11} \in \mathcal{OC}^{2\bar{k}\times 2\bar{k}}$ and $\boldsymbol{\Psi}_{21} \in \mathcal{OC}^{2\bar{k}\times 2\bar{k}}$ are independent of $\boldsymbol{H}_{11}$ and $\boldsymbol{H}_{21}$, respectively. Therefore, the matrices $\overline{\boldsymbol{H}}_{11}$ and $\overline{\boldsymbol{H}}_{21}$, defined in (42), have Rayleigh distribution. Similar arguments are valid for $\overline{\boldsymbol{H}}_{12}$ and $\overline{\boldsymbol{H}}_{22}$. Therefore, the system is decomposed into two broadcast subchannels, each with the Rayleigh distribution. Hence, the sum–rate of the MIMO broadcast subchannel, viewed from transmitter t, is approximated by [20]

$$2\bar{k}[\log_2(P_t) - \mathcal{L}_\infty(2\bar{k}, 2\bar{k})] + o(1). \tag{51}$$

By summation of the approximated formulas for the two MIMO broadcast subchannels, (48) is obtained.

In [19], it is proven that the power offset for a MIMO Rayleigh channel with M transmit and N receive antennas is obtained by (49). By substituting $M = 4\bar{k}$ and $N = 6\bar{k}$ in (49), and subtracting (49) from (48),(50) is derived. □

V. Joint Design

The decomposition schemes proposed in Section III simplify the signaling design and the performance evaluation for the X channels. However, such decomposition schemes deteriorate the performance of the system because: i) $\boldsymbol{\Psi}_{rt}$, $r,t = 1,2$, are chosen such that the interference terms are forced to be zero, while the statistical properties of the interference should be exploited to design these filters, ii) $\boldsymbol{Q}_t$, $t = 1,2$ are chosen randomly, while the gain of the channel matrices in the different directions should be considered in choosing $\boldsymbol{Q}_t$, $t = 1,2$. For example, consider an X channel with $(2,2,3,3)$ antennas. In Section III-A, the receive filters $\boldsymbol{\Psi}_{rt}$, $r,t = 1,2$, are chosen such that the interference of each broadcast subchannel over the other one is forced to be zero. In low-SNR regimes, the performance of the system is improved by choosing whitening filters for $\boldsymbol{\Psi}_{rt}, r,t = 1,2$, instead of zero-forcing filters. In high SNR, the whitening filters converge to zero-forcing filters, and the resulting improvement diminishes. Note that in the X channel with $(2,2,3,3)$, the transmit filters $\boldsymbol{Q}_t$, $t = 1,2$, are such that the entire two-dimensional spaces available at transmitters one and two are used for signaling. Therefore, we cannot improve the signaling scheme by modifying $\boldsymbol{Q}_t, t = 1,2$.

In a system with $(3,3,3,3)$ antennas, the same arguments for $\boldsymbol{\Psi}_{rt}, r,t = 1,2$ are valid. In this case, the transmit filters $\boldsymbol{Q}_t, t = 1,2$, are chosen randomly, therefore the signaling space at each transmitter is confined to a randomly selected two-dimensional subspace of a three-dimensional space. One can take advantage of the degrees of freedom available for choosing $\boldsymbol{Q}_t$ to find the signaling subspaces at transmitters one and two for which the channels offer the highest gains.

Optimizing the filters $\boldsymbol{Q}_t$ and $\boldsymbol{\Psi}_{rt}, r,t = 1,2$, depends on the signaling scheme employed for the MIMO broadcast or multiple-access subchannels. On the other hand, designing the signaling schemes for the subchannels depends on the selected filters. Therefore, we have to jointly develop the design parameters. In what follows, we elaborate a joint design scheme based on a generalized version of zero-forcing dirty paper coding (ZF-DPC) scheme, presented in [21], for the resulting broadcast subchannels in Scheme I. In this scheme, the number of data streams $\mu_{rt}, r,t = 1,2$, and also integer parameters $\mu'_{rt}, r,t = 1,2$ are selected as explained in Section III-A. In addition, we use filters $\boldsymbol{Q}_t$ and $\boldsymbol{\Psi}^\dagger_{rt}$, r,t, in a similar fashion as shown in Fig. 1, but with a different design.

According to the generalized ZF-DPC, explained in [21] for MIMO broadcast channels, the vectors $\tilde{\boldsymbol{s}}_t, t = 1,2$, are equal to linear superpositions of some modulation vectors where the data is embedded in the coefficients. The modulation matrix $\boldsymbol{V}_t \in \mathcal{OC}^{(\mu_{1t}+\mu_{2t})\times(\mu_{1t}+\mu_{2t})}$ is defined as

$$\boldsymbol{V}_t = \left[\boldsymbol{v}_t^{(1)}, \boldsymbol{v}_t^{(2)}, \ldots, \boldsymbol{v}_t^{(\mu_{1t}+\mu_{2t})}\right] \tag{52}$$

where $\boldsymbol{v}_t^{(i)}, i = 1,\ldots,\mu_{1t}+\mu_{2t}$, denote the modulation vectors, employed by transmitter t, to send μ_{1t} data streams to receiver one and μ_{2t} data streams to receiver two. The vectors $\tilde{\boldsymbol{s}}_1$ and $\tilde{\boldsymbol{s}}_2$ are equal to

$$\tilde{\boldsymbol{s}}_1 = \boldsymbol{V}_1 \boldsymbol{d}_1 \tag{53}$$

$$\tilde{\boldsymbol{s}}_2 = \boldsymbol{V}_2 \boldsymbol{d}_2 \tag{54}$$

where the vector $\boldsymbol{d}_t \in \mathcal{C}^{(\mu_{1t}+\mu_{2t})\times 1}$ represents the $\mu_{1t} + \mu_{2t}$ streams of independent data. The covariance of the vector $\boldsymbol{d}_t$ is denoted by the diagonal matrix $\boldsymbol{P}_t$, i.e., $E[\boldsymbol{d}_t\boldsymbol{d}_t^\dagger] = \boldsymbol{P}_t, t = 1,2$. At transmitter t, the data streams which modulate the vectors

$\boldsymbol{v}_t^{(i)}$, $i=1,\ldots,\mu'_{1t}$, and $i=\mu'_{1t}+\mu'_{21}+1,\ldots,\mu_{1t}+\mu'_{2t}$, are intended for the receiver one, and the data streams which modulate the vectors $\boldsymbol{v}_t^{(i)}$, $i=\mu'_{1t}+1,\ldots,\mu'_{1t}+\mu'_{2t}$, and $i=\mu_{1t}+\mu'_{2t}+1,\ldots,\mu_{1t}+\mu_{2t}$, are intended for receiver two. We define $\boldsymbol{d}_{1t}$ and $\boldsymbol{d}_{2t}$ as

$$\boldsymbol{d}_{1t}=\begin{bmatrix}\boldsymbol{d}_t(1:\mu'_{11})\\ \boldsymbol{d}_t(\mu'_{11}+\mu'_{21}+1:\mu_{11}+\mu'_{21})\end{bmatrix} \tag{55}$$

and

$$\boldsymbol{d}_{2t}=\begin{bmatrix}\boldsymbol{d}_t(\mu'_{11}+1:\mu'_{11}+\mu'_{21})\\ \boldsymbol{d}_t(\mu_{11}+\mu'_{21}+1:\mu_{11}+\mu_{21})\end{bmatrix} \tag{56}$$

which represent the data streams, sent by transmitter t to receivers one and two, respectively. The modulation and demodulation vectors are designed such that the data stream i has no interference over the data stream j for $j<i$. Choosing the codeword for the data stream j, the interference of the data stream j over data stream i is noncausally known, and therefore can be effectively canceled out based on the dirty-paper coding (DPC) theorem [22]. However, if the data streams i and j are sent to the same receiver, none of them has interference over the other, and DPC is not needed. At receiver one, to decode $\boldsymbol{d}_{11}$, the signal coming from transmitter two, i.e., $\widetilde{\boldsymbol{H}}_{12}\boldsymbol{V}_{12}\boldsymbol{d}_2$, is treated as interference, therefore, the covariance of the interference plus noise $\boldsymbol{R}_{11}$ is equal to

$$\boldsymbol{R}_{11}=\widetilde{\boldsymbol{H}}_{12}\boldsymbol{V}_2\boldsymbol{P}_2\boldsymbol{V}_2^{\dagger}\widetilde{\boldsymbol{H}}_{12}+\boldsymbol{I} \tag{57}$$

where $\widetilde{\boldsymbol{H}}_{12}$ is defined in (107). The received vector $\boldsymbol{y}_1$ is passed through the whitening filter $\boldsymbol{\Psi}_{11}^{\dagger}=\boldsymbol{R}_{11}^{-\frac{1}{2}}$. The output of $\boldsymbol{\Psi}_{11}^{\dagger}$ is passed through the filter $\boldsymbol{U}_{11}^{\dagger}$ which maximizes the effective SNR. The design of $\boldsymbol{U}_{rt}^{\dagger}$, $r,t=1,2$, is explained later. Similarly, to decode $\boldsymbol{d}_{21}$ at receiver two, the signal from transmitter two, i.e., $\widetilde{\boldsymbol{H}}_{22}\boldsymbol{V}_2\boldsymbol{d}_2$ is treated as interference. The received vector $\boldsymbol{y}_2$ is passed through the whitening filter $\boldsymbol{\Psi}_{21}^{\dagger}=\boldsymbol{R}_{21}^{-\frac{1}{2}}$, where

$$\boldsymbol{R}_{21}=\widetilde{\boldsymbol{H}}_{22}\boldsymbol{V}_2\boldsymbol{P}_2\boldsymbol{V}_2^{\dagger}\widetilde{\boldsymbol{H}}_{22}+\boldsymbol{I}. \tag{58}$$

The output of $\boldsymbol{\Psi}_{21}^{\dagger}$ is passed through the filter $\boldsymbol{U}_{21}^{\dagger}$ which maximizes the effective SNR.

Let us assume that the modulation matrix $\boldsymbol{V}_2$, the covariance matrix $\boldsymbol{P}_2$, and the transmit filter $\boldsymbol{Q}_2$ are known, therefore, one can compute $\boldsymbol{\Psi}_{11}^{\dagger}$ and $\boldsymbol{\Psi}_{21}^{\dagger}$. In the sequel, we explain how to choose $\boldsymbol{Q}_1$, $\boldsymbol{V}_1$, $\boldsymbol{P}_1$, $\boldsymbol{U}_{11}$, and $\boldsymbol{U}_{21}$.

The following algorithm is proposed to compute the columns of the matrix $\boldsymbol{Q}_1\in\mathcal{OC}^{M_1\times(\mu_{11}+\mu_{21})}$. The proposed algorithm is greedy in the sense that in each step, the direction along which the corresponding link has the highest gain is added to the columns of the matrix $\boldsymbol{Q}_1$. In the algorithm, the following four sets of vectors are sequentially included in the columns of $\boldsymbol{Q}_1$: i) the μ'_{11} mutually orthogonal directions for which the equivalent channel matrix $\boldsymbol{\Psi}_{11}^{\dagger}\boldsymbol{H}_{11}$ has the highest gains, ii) the μ'_{21} mutually orthogonal directions for which the equivalent channel matrix $\boldsymbol{\Psi}_{21}^{\dagger}\boldsymbol{H}_{21}$ has the highest gains, iii) if $\mu_{11}-\mu'_{11}\neq 0$, a set of directions such that $\mathrm{N}(\boldsymbol{H}_{21})\in\Omega(\boldsymbol{Q}_1)$, iv) if $\mu_{21}-\mu'_{21}\neq 0$, a set of directions such that $\mathrm{N}(\boldsymbol{H}_{11})\in\Omega(\boldsymbol{Q}_1)$. Each set of vectors are chosen orthogonal to the previously selected columns. In what follows, we detail the algorithm in four stages.

Stage I

- Choose $\boldsymbol{q}_1^{(i)}$, $i=1,\ldots,\mu'_{11}$, as μ'_{11} right singular vectors (RSV) corresponding to the μ'_{11} largest singular values of the matrix $\boldsymbol{\Psi}_{11}^{\dagger}\boldsymbol{H}_{11}$.

Stage II

- Choose $\boldsymbol{\Phi}_1=[\boldsymbol{\phi}_1,\ldots,\boldsymbol{\phi}_{\mu_{11}+\mu_{21}-\mu'_{11}}]$ such that $[\boldsymbol{\Phi}_1,\boldsymbol{q}_1^{(1)},\ldots,\boldsymbol{q}_1^{(\mu'_{11})}]$ forms a unitary matrix.
- Choose $\boldsymbol{q'}_1^{(i)}$, $i=1,\ldots,\mu'_{21}$, as the μ'_{21} RSVs corresponding to the μ'_{21} largest singular values of the matrix $\boldsymbol{\Psi}_{21}^{\dagger}\boldsymbol{H}_{21}\boldsymbol{\Phi}_1$.
- Let $\boldsymbol{q}_1^{(\mu'_{11}+i)}=\boldsymbol{\Phi}_1\boldsymbol{q'}_1^{(i)}$, $i=1,\ldots,\mu'_{21}$.

Stage III

- If $\mu_{11}-\mu'_{11}\neq 0$, then choose $\boldsymbol{q}_1^{(i)}$, $i=\mu'_{11}+\mu'_{21}+1,\ldots,\mu_{11}+\mu'_{21}$, such that
$$\Omega\left(\left[\boldsymbol{q}_1^{(1)},\ldots,\boldsymbol{q}_1^{(\mu_{11}+\mu'_{21})}\right]\right)=\Omega\left(\left[\boldsymbol{q}_1^{(1)},\ldots,\boldsymbol{q}_1^{(\mu'_{11}+\mu'_{21})},\mathrm{N}(\boldsymbol{H}_{21})\right]\right).$$

Stage IV

- If $\mu_{21}-\mu'_{21}\neq 0$, then choose $\boldsymbol{q}_1^{(i)}$, $i=\mu_{11}+\mu'_{21}+1,\ldots,\mu_{11}+\mu_{21}$, such that
$$\Omega\left(\left[\boldsymbol{q}_1^{(1)},\ldots,\boldsymbol{q}_1^{(\mu_{11}+\mu_{21})}\right]\right)=\Omega\left(\left[\boldsymbol{q}_1^{(1)},\ldots,\boldsymbol{q}_1^{(\mu_{11}+\mu'_{21})},\mathrm{N}(\boldsymbol{H}_{11})\right]\right).$$

The intuition behind this order of dimension selection is as follows. Consider transmitter one. μ_{11} and μ_{21} space dimensions are needed to send μ_{11} and μ_{21} data streams to receivers one and two, respectively. However, $\mu_{11}-\mu'_{11}$ out of μ_{11} dimensions and $\mu_{21}-\mu'_{21}$ out of μ_{21} dimensions have to be chosen from $\mathrm{N}(\boldsymbol{H}_{21})$ and $\mathrm{N}(\boldsymbol{H}_{11})$, respectively. In this algorithm, μ'_{11} and μ'_{21} dimensions are first chosen in a recursive and greedy manner to exploit the highest gains through the channel matrices $\boldsymbol{H}_{11}$ and $\boldsymbol{H}_{21}$, respectively. Then, the transmit space is extended to include $\mu_{11}-\mu'_{11}$ and $\mu_{21}-\mu'_{21}$ dimensions from $\mathrm{N}(\boldsymbol{H}_{21})$ and $\mathrm{N}(\boldsymbol{H}_{11})$, respectively. Apparently, if this order is changed and the dimensions are selected initially from $\mathrm{N}(\boldsymbol{H}_{21})$ and $\mathrm{N}(\boldsymbol{H}_{11})$, then we may lose the dimensions which provide highest gains through $\boldsymbol{H}_{11}$ and $\boldsymbol{H}_{21}$.

After computing $\boldsymbol{Q}_1$, the broadcast subchannel with $\overline{\boldsymbol{H}}_{11}$ and $\overline{\boldsymbol{H}}_{21}$, defined in Section III-A as $\overline{\boldsymbol{H}}_{r1}=\boldsymbol{\Psi}_{r1}^{\dagger}\boldsymbol{H}_{r1}\boldsymbol{Q}_1$, $r=1,2$, is formed. Here, we explain how to choose the modulation and demodulation vectors for this broadcast subchannel, based on the scheme presented in [21]. In the scheme presented in [21], the modulation vectors for different users can be selected iteratively in a specific order. Here, the modulation vectors are selected in the following order: i) μ'_{11} modulation vectors for receiver one, ii) μ'_{21} modulation vectors for receiver two, iii) $\mu_{11}-\mu'_{11}$ modulation vectors for receiver one, iv) $\mu_{21}-\mu'_{21}$ modulation vectors for receiver two. Here is the detail of the proposed scheme to find the modulation and the demodulation vectors.

Step one–Choosing μ'_{11} modulation vectors for receiver one

1) Respectively choose $\boldsymbol{v}_1^{(i)}$ and $\boldsymbol{u}_{11}^{(i)}$, $i=1,\dots,\mu'_{11}$, as RSV and left singular vector (LSV), corresponding to the ith largest singular value $\sigma_{11}^{(i)}$ of the matrix $\overline{\boldsymbol{H}}_{11}$. Therefore, we have [23]

$$\sigma_{11}^{(i)} = \|\overline{\boldsymbol{H}}_{11}\boldsymbol{v}_1^{(i)}\|, \quad i=1,\dots,\mu'_{11} \tag{59}$$

$$\boldsymbol{u}_{11}^{(i)} = \frac{\overline{\boldsymbol{H}}_{11}\boldsymbol{v}_{11}^{(i)}}{\sigma_{11}^{(i)}}, \quad i=1,\dots,\mu'_{11}. \tag{60}$$

With the above choice of the matrix $\boldsymbol{Q}_1$, it is easy to see that $\boldsymbol{v}_1^{(i)}$ is equal to the column i of the identity matrix $\boldsymbol{I}_{(\mu_{11}+\mu_{21})\times(\mu_{11}+\mu_{21})}$, for $i=1,\dots,\mu'_{11}$.

Step two–Choosing μ'_{21} modulation vectors for receiver two

1) Define $\boldsymbol{\varphi}_1^{(1)}, \dots, \boldsymbol{\varphi}_1^{(\mu_{11}+\mu_{21}-\mu'_{11})}$ such that $[\boldsymbol{v}_1^{(1)},\dots,\boldsymbol{v}_1^{(\mu'_{11})},\boldsymbol{\varphi}_1^{(1)},\dots,\boldsymbol{\varphi}_1^{(\mu_{11}+\mu_{21}-\mu'_{11})}]$ forms a unitary matrix. Then, define $\widehat{\overline{\boldsymbol{H}}}_{21}$ as

$$\widehat{\overline{\boldsymbol{H}}}_{21} = \overline{\boldsymbol{H}}_{21}\left[\boldsymbol{\varphi}_1^{(1)},\dots,\boldsymbol{\varphi}_1^{(\mu_{11}+\mu_{21}-\mu'_{11})}\right]. \tag{61}$$

2) Respectively choose $\overline{\boldsymbol{v}}_{21}^{(i)}$ and $\boldsymbol{u}_{21}^{(i)}$ as the RSV and LSV, corresponding to the ith largest singular value $\sigma_{21}^{(i)}$ of the matrix $\widehat{\overline{\boldsymbol{H}}}_{21}$. Therefore, we have

$$\sigma_{21}^{(i)} = \|\widehat{\overline{\boldsymbol{H}}}_{21}\overline{\boldsymbol{v}}_{21}^{(i)}\|, \quad i=1,\dots,\mu'_{21} \tag{62}$$

$$\boldsymbol{u}_{21}^{(i)} = \frac{\widehat{\overline{\boldsymbol{H}}}_{21}\overline{\boldsymbol{v}}_{21}^{(i)}}{\sigma_{21}^{(i)}}, \quad i=1,\dots,\mu'_{21}. \tag{63}$$

Then, let

$$\boldsymbol{v}_1^{(\mu'_{11}+i)} = \left[\boldsymbol{\varphi}_1^{(1)},\dots,\boldsymbol{\varphi}_1^{(\mu_{11}+\mu_{21}-\mu'_{11})}\right]\overline{\boldsymbol{v}}_{21}^{(i)}, \quad i=1,\dots,\mu'_{21}. \tag{64}$$

It is easy to see that with the aforementioned choice of $\boldsymbol{Q}_1$, $\boldsymbol{v}_1^{(\mu'_{11}+i)}$ is equal to the column $\mu'_{11}+i$ of the matrix $\boldsymbol{I}_{(\mu_{11}+\mu_{21})\times(\mu_{11}+\mu_{21})}$, for $i=1,\dots,\mu'_{21}$.

Step three–Choosing $\mu_{11}-\mu'_{11}$ modulation vectors for receiver one

1) Define $\boldsymbol{\varphi}_2^{(1)}, \dots, \boldsymbol{\varphi}_2^{(\mu_{11}+\mu_{21}-\mu'_{11}-\mu'_{21})}$ such that $\left[\boldsymbol{v}_1^{(1)},\dots,\boldsymbol{v}_1^{(\mu'_{11}+\mu'_{21})},\boldsymbol{\varphi}_2^{(1)},\dots,\boldsymbol{\varphi}_2^{(\mu_{11}+\mu_{21}-\mu'_{11}-\mu'_{21})}\right]$ forms a unitary matrix. Then, define $\widehat{\overline{\boldsymbol{H}}}_{11}$ as

$$\widehat{\overline{\boldsymbol{H}}}_{11} = \overline{\boldsymbol{H}}_{11}\left[\boldsymbol{\varphi}_2^{(1)},\dots,\boldsymbol{\varphi}_2^{(\mu_{11}+\mu_{21}-\mu'_{11}-\mu'_{21})}\right]. \tag{65}$$

2) Respectively choose $\overline{\boldsymbol{v}}_{11}^{(i)}$ and $\boldsymbol{u}_{11}^{(i+\mu'_{11})}$ as the RSV and LSV, corresponding to the i^{th} largest singular value of the matrix $\widehat{\overline{\boldsymbol{H}}}_{11}$, denoted by $\sigma_{11}^{(i+\mu'_{11})}$, for $i=1,\dots,\mu_{11}-\mu'_{11}$. Therefore, we have

$$\sigma_{11}^{(i+\mu'_{11})} = \|\widehat{\overline{\boldsymbol{H}}}_{11}\overline{\boldsymbol{v}}_{11}^{(i)}\|, \quad i=1,\dots,\mu_{11}-\mu'_{11} \tag{66}$$

$$\boldsymbol{u}_{11}^{(i+\mu'_{11})} = \frac{\widehat{\overline{\boldsymbol{H}}}_{11}\overline{\boldsymbol{v}}_{11}^{(i)}}{\sigma_{11}^{(i+\mu'_{11})}}, \quad i=1,\dots,\mu_{11}-\mu'_{11}. \tag{67}$$

Then

$$\boldsymbol{v}_1^{(\mu'_{11}+\mu'_{21}+i)} = \left[\boldsymbol{\varphi}_2^{(1)},\dots,\boldsymbol{\varphi}_2^{(\mu_{11}+\mu_{21}-\mu'_{11}-\mu'_{21})}\right]\overline{\boldsymbol{v}}_{11}^{(i)}, \quad i=1,\dots,\mu_{11}-\mu'_{11}. \tag{68}$$

Step four–Choosing $\mu_{21}-\mu'_{21}$ modulation vectors for receiver two

1) Define $\boldsymbol{\varphi}_3^{(1)}, \dots, \boldsymbol{\varphi}_3^{(\mu_{21}-\mu'_{21})}$ such that $[\boldsymbol{v}_1^{(1)},\dots,\boldsymbol{v}_1^{(\mu_{11}+\mu'_{21})},\boldsymbol{\varphi}_3^{(1)},\dots,\boldsymbol{\varphi}_3^{(\mu_{21}-\mu'_{21})}]$ forms a unitary matrix. Then, define $\widehat{\widehat{\overline{\boldsymbol{H}}}}_{21}$ as

$$\widehat{\widehat{\overline{\boldsymbol{H}}}}_{21} = \overline{\boldsymbol{H}}_{21}\left[\boldsymbol{\varphi}_3^{(1)},\dots,\boldsymbol{\varphi}_3^{(\mu_{21}-\mu'_{21})}\right]. \tag{69}$$

2) Respectively choose $\overline{\overline{\boldsymbol{v}}}_{21}^{(i)}$ and $\boldsymbol{u}_{21}^{(i+\mu'_{21})}$ as RSV and LSV, corresponding to the ith largest singular value of the matrix $\widehat{\widehat{\overline{\boldsymbol{H}}}}_{11}$, denoted by $\sigma_{21}^{(i+\mu'_{21})}$, for $i=1,\dots,\mu_{21}-\mu'_{21}$. Therefore, we have

$$\sigma_{21}^{(i+\mu'_{21})} = \|\widehat{\widehat{\overline{\boldsymbol{H}}}}_{21}\overline{\overline{\boldsymbol{v}}}_{21}^{(i)}\|, \quad i=1,\dots,\mu_{21}-\mu'_{21} \tag{70}$$

$$\boldsymbol{u}_{21}^{(i+\mu'_{21})} = \frac{\widehat{\overline{\boldsymbol{H}}}_{21}\overline{\overline{\boldsymbol{v}}}_{21}^{(i)}}{\sigma_{21}^{(i+\mu'_{21})}}, \quad i=1,\dots,\mu_{21}-\mu'_{21}. \tag{71}$$

Then, let

$$\boldsymbol{v}_1^{(\mu_{11}+\mu'_{21}+i)} = \left[\boldsymbol{\varphi}_3^{(1)},\dots,\boldsymbol{\varphi}_3^{(\mu_{21}-\mu'_{21})}\right]\overline{\overline{\boldsymbol{v}}}_{21}^{(i)}, \quad i=1,\dots,\mu_{11}-\mu'_{11}. \tag{72}$$

As shown in [21], by using this scheme, the broadcast channel, viewed from transmitter one is reduced to a set of parallel channels with gains $\sigma_{11}^{(i)}$, $i=1,\dots,\mu_{11}$ and $\sigma_{21}^{(j)}$, $j=1,\dots,\mu_{21}$. For power allocation, the power P_1 can be equally divided among the data streams or the water-filling algorithm can be used for optimal power allocation [24].

Similar procedure is applied for transmitter two to compute $\boldsymbol{Q}_2, \boldsymbol{V}_2, \boldsymbol{U}_{12}, \boldsymbol{U}_{22}, \boldsymbol{P}_2$, where

$$\boldsymbol{R}_{22} = \boldsymbol{H}_{21}\boldsymbol{V}_1\boldsymbol{P}_1\boldsymbol{V}_1^\dagger\boldsymbol{H}_{21}^\dagger + \boldsymbol{I} \tag{73}$$

$$\boldsymbol{\Psi}_{22}^\dagger = \boldsymbol{R}_{22}^{-\frac{1}{2}} \tag{74}$$

$$\boldsymbol{R}_{12} = \boldsymbol{H}_{11}\boldsymbol{V}_1\boldsymbol{P}_1\boldsymbol{V}_1^\dagger\boldsymbol{H}_{11}^\dagger + \boldsymbol{I} \tag{75}$$

$$\boldsymbol{\Psi}_{12}^\dagger = \boldsymbol{R}_{12}^{-\frac{1}{2}}. \tag{76}$$

Note that to compute $\boldsymbol{Q}_1$, $\boldsymbol{V}_1$, and $\boldsymbol{P}_1$, we need to know $\boldsymbol{Q}_2$, $\boldsymbol{V}_2$, and $\boldsymbol{P}_2$ ($\boldsymbol{\Psi}_{11}$, and $\boldsymbol{\Psi}_{21}$ are functions of $\boldsymbol{Q}_2$, $\boldsymbol{V}_2$, and $\boldsymbol{P}_2$), and *vice versa*. To derive the modulation vectors, we can randomly initialize the matrices, and iteratively follow the scheme, until the resulting matrices converge. Simulation results show that the algorithm converges very fast.

It is possible to improve the decomposition Scheme II, presented in Appendix I, by jointly design the filters.

VI. Simulation Results

In the simulation part, we assume that the entries of the channel matrices have complex normal distribution with zero mean and unit variance.

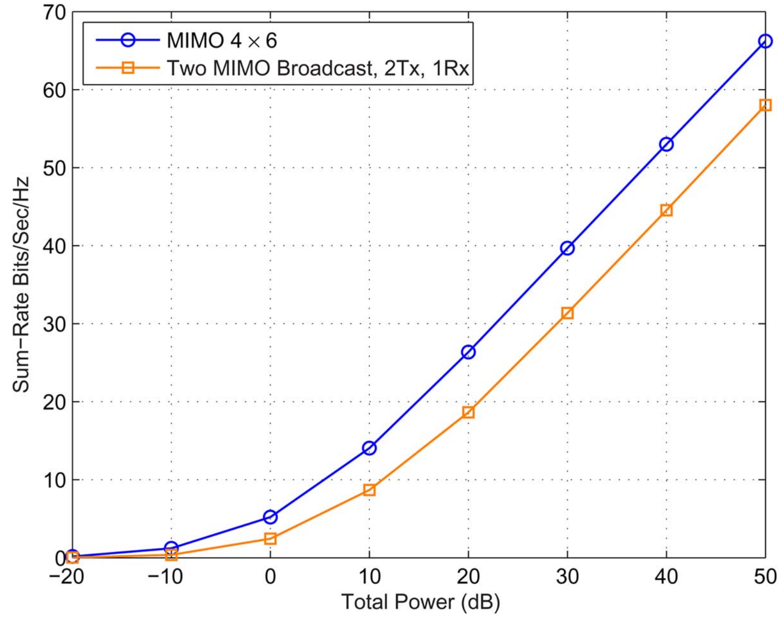

Fig. 6. The sum–capacity of point-to-point MIMO channel with four transmit and six receive antennas, and the sum–rate of the X channel with $(2,2,3,3)$ antennas achieved based on decomposition Scheme I.

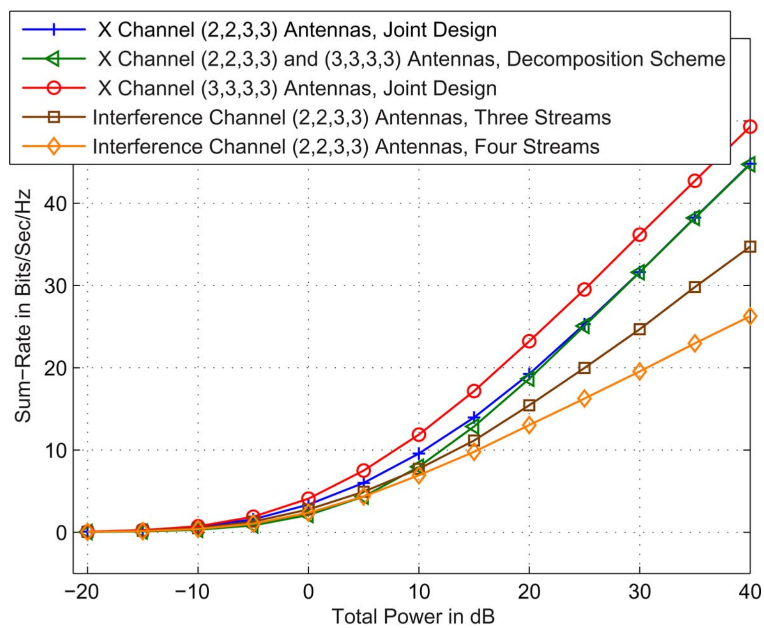

Fig. 7. The sum–rate of the X channels using ZF-DPC scheme over the decomposed channels, the sum–rate of the X channels achieved by jointly designed ZF-DPC scheme, and the sum–rate of the interference channel with three and four data streams.

Fig. 6 shows the sum–rate versus power for a X channel with $(2,2,3,3)$ antennas, where the decomposition scheme presented in Section III is employed. Therefore, the achievable sum–rate is indeed equal to twice of the sum–capacity of a MIMO broadcast channel with two transmit antennas, and two users each with one antenna. The sum–capacity of the MIMO broadcast channel is fully characterized in [25]–[27]. To compute the sum–capacity, the effective algorithm presented in [28] is utilized. As a comparison, the capacity of a point-to-point MIMO channel with four transmit and six receive antennas is depicted. It is easy to see that both curves have the same slope (multiplexing gain). In addition, as expected by (50), the sum–rate of the X channel has 6.2-dB power loss in comparison with that of the MIMO channel.

Fig. 7 shows the sum–rate versus power for an X channel with $(2,2,3,3)$ and $(3,3,3,3)$ antennas, where ZD-DPC scheme is used. As it is shown in Fig. 7, for the case of $(2,2,3,3)$ antennas, the joint design scheme has better performance than the decomposition scheme in low-SNR regimes. The improvement is mainly due to utilizing whitening filters instead of zero-forcing filters. It is easy to see that in the high SNR, the whitening filters converge to zero-forcing filters. Note that in this case, optimizing $\boldsymbol{Q}_t$, $t=1,2$ offers no improvement. The reason is that the entire two-dimensional space available at each transmitter is utilized and there is no room for improvement. As depicted in Fig. 7, for the case of $(3,3,3,3)$-antenna X channel, the joint design scheme has better performance as compared with the decomposition scheme in both high- and low-SNR regimes. The improvement relies on the fact that in this case at each transmitter, a two-dimensional subspace of the three-dimensional space is needed for signaling. By using the scheme presented in Section V, a subspace for which the channel gains are optimal is chosen. Note that the sum–rate of the $(3,3,3,3)$-antenna X channel, where the decomposition scheme is applied, is the same as that of $(2,2,3,3)$-antenna X channel. The reason is that in this case, the matrices $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ are randomly selected from the set of 3×2 unitary matrices. Therefore, the resulting system is equivalent to a $(2,2,3,3)$-antenna X channel.

Fig. 7 also shows the curves of the sum–rate versus the total power for $(2,2,3,3)$-antenna interference channels for two signaling scenarios. In the first scenario, each transmitter sends two data streams to the corresponding receiver. Each receiver employs optimal whitening filter, treating the signal coming from the other transmitter as interference. In this case, there is only one interference-free dimension available at each receiver. Therefore, the rate of one of the data streams sent by each transmitter linearly increases with the logarithm of the total power, while the rate of the other data stream converges to a constant number. Therefore, this scheme achieves the overall MG of two. In the second scenario, transmitters one and two send two and one data streams, respectively. $2/3$ of the total power is allocated to transmitter one and the rest is allocated to transmitter two. In this case, two of three space dimensions at receiver one is available for the signal coming from transmitter one. In addition, one out of three space dimensions at receiver two is available for the signal coming from transmitter two. Therefore, this scheme achieves the overall MG of three. As mentioned earlier, this is the optimal MG for $(2,2,3,3)$-antenna interference channel, which is less than achievable MG by the corresponding X channel.

VII. Conclusion

In a multiple-antenna system with two transmitters and two receivers, a new noncooperative scenario of data communication is studied in which each receiver receives data from both transmitters. It is shown that by using some linear filters at the transmitters and at the receivers, the system is decomposed into two broadcast or two multiple-access subchannels. Using the decomposition scheme, it is shown that this signaling method outperforms other known noncooperative schemes in terms of the achieved multiplexing gain. In particular, it is shown that for a system with $\left(\left\lceil\frac{1}{2}\left\lfloor\frac{4N}{3}\right\rfloor\right\rceil,\left\lfloor\frac{1}{2}\left\lfloor\frac{4N}{3}\right\rfloor\right\rfloor,N,N\right)$ and $\left(N,N,\left\lceil\frac{1}{2}\left\lfloor\frac{4N}{3}\right\rfloor\right\rceil,\left\lfloor\frac{1}{2}\left\lfloor\frac{4N}{3}\right\rfloor\right\rfloor\right)$ antennas, the multiplexing gain of $\left\lfloor\frac{4N}{3}\right\rfloor$ is achievable, which is the MG of the system where full cooperation between the transmitters or between the receivers is provided.

Appendix I
Scheme II–Decomposition of the System Into Two Multiple-Access Subchannels

This scheme is indeed the dual of the scheme one, detailed in Section III-A. As depicted in Fig. 3, in this scheme, the parallel transmit filters $\boldsymbol{\Psi}_{11}$ and $\boldsymbol{\Psi}_{21}$ are employed at transmitter one, and the parallel transmit filters $\boldsymbol{\Psi}_{12}$ and $\boldsymbol{\Psi}_{22}$ are employed at transmitter two. Therefore, the transmitted vectors are equal to

$$\boldsymbol{s}_1 = \boldsymbol{\Psi}_{11}\boldsymbol{s}_{11} + \boldsymbol{\Psi}_{21}\boldsymbol{s}_{21} \tag{77}$$
$$\boldsymbol{s}_2 = \boldsymbol{\Psi}_{12}\boldsymbol{s}_{12} + \boldsymbol{\Psi}_{22}\boldsymbol{s}_{22} \tag{78}$$

where $\boldsymbol{s}_{rt} \in \mathcal{C}^{\mu_{rt}\times 1}$ contains μ_{rt} data streams from transmitter t intended to receiver r. The transmit filter $\boldsymbol{\Psi}_{11}$ nulls out the interference of the μ_{11} data streams, sent from transmitter one to receiver one, at receiver two. Similarly, the transmit filter $\boldsymbol{\Psi}_{21}$ nulls out the interference of the μ_{21} data streams sent from transmitter one to receiver two at receiver one. In a similar fashion, at transmitter two, the two parallel transmit filters $\boldsymbol{\Psi}_{22}$ and $\boldsymbol{\Psi}_{12}$ are employed.

At receiver r terminal, the received vector is passed through the receive filter $\boldsymbol{Q}_r^{\dagger}$, where $\boldsymbol{Q}_r \in \mathcal{OC}^{N_r\times(\mu_{r1}+\mu_{r2})}$

$$\tilde{\boldsymbol{y}}_r = \boldsymbol{Q}_r^{\dagger}\boldsymbol{y}_r, \quad r = 1,2. \tag{79}$$

The functionalities of the receive filters $\boldsymbol{Q}_t$, $t = 1,2$, are i) to map the received signal in a $(\mu_{r1}+\mu_{r2})$-dimensional subspace, which allows us to null out the interference terms by using transmit filters $\boldsymbol{\Psi}_{rt}$, $r,t = 1,2$, and ii) to exploit the null spaces of the channel matrices to attain the highest MG.

Similar to the previous section, it is shown that if the numbers of data streams μ_{rt}, $r,t = 1,2$, satisfy a set of inequalities, then it is possible to deign $\boldsymbol{Q}_t$ and $\boldsymbol{\Psi}_{rt}$ to meet the desired features explained earlier. Consequently, the system is decomposed into two noninterfering MIMO multiple-access subchannels (see Fig. 4).

Next, we explain how to select the design parameters including the numbers of data streams μ_{rt}, $r,t = 1,2$ and the transmit/receive filters. Again, the primary objective is to prevent the saturation of the rate of each stream in the high-SNR regime. In other words, the MG of the system is $\mu_{11}+\mu_{12}+\mu_{21}+\mu_{22}$.

Similar to the previous subsection, we define the parameters ζ_{rt}, $r,t = 1,2$, as follows:

- ζ_{11} denotes the dimension of $\Omega(\boldsymbol{H}_{21}^{\dagger}\boldsymbol{Q}_2)$;
- ζ_{21} denotes the dimension of $\Omega(\boldsymbol{H}_{11}^{\dagger}\boldsymbol{Q}_1)$;
- ζ_{12} denotes the dimension of $\Omega(\boldsymbol{H}_{22}^{\dagger}\boldsymbol{Q}_2)$;
- ζ_{22} denotes the dimension of $\Omega(\boldsymbol{H}_{12}^{\dagger}\boldsymbol{Q}_1)$.

To facilitate the derivations, we use the auxiliary variables M'_t, N'_r, and μ'_{rt}, for $r,t = 1,2$. For each case, the variables M'_t and N'_r are computed directly as a function of M_t and N_r for $r,t = 1,2$. Then, the auxiliary integer variables μ'_{rt}, $r,t = 1,2$, are selected such that the following constraints are satisfied:

$$\mu'_{11}: \quad \mu'_{11}+\mu'_{21}+\mu'_{22} \le M'_1, \tag{80}$$
$$\mu'_{21}: \quad \mu'_{11}+\mu'_{21}+\mu'_{12} \le M'_1 \tag{81}$$
$$\mu'_{22}: \quad \mu'_{22}+\mu'_{12}+\mu'_{11} \le M'_2 \tag{82}$$
$$\mu'_{12}: \quad \mu'_{22}+\mu'_{12}+\mu'_{21} \le M'_2 \tag{83}$$
$$\mu'_{11}+\mu'_{12} \le N'_1 \tag{84}$$
$$\mu'_{22}+\mu'_{21} \le N'_2. \tag{85}$$

Each of the first four inequalities corresponds to one of the parameters μ'_{rt}, $r,t=1,2$, in the sense that if μ'_{rt}, $r,t=1,2$, is zero, the corresponding inequality is removed from the set of constraints. After choosing μ'_{rt}, $r,t=1,2$, for each case, μ_{rt}, $r,t=1,2$, are computed as function of μ'_{rt}, $r,t=1,2$.

Similar to Scheme I, to achieve the highest MG, we choose $\mu'_{rt}, r,t=1,2$ subject to (80)–(85) such that $\mu'_{11}+\mu'_{12}+\mu'_{21}+\mu'_{22}$ is maximum.

Next, for each of the four cases, we explain the following.

i) How to compute the auxiliary variables M'_t and N'_r as a function of M_t and N_r, $r,t=1,2$.
ii) After choosing the auxiliary variables μ'_{rt}, $r,t=1,2$, satisfying (80)–(85), how to compute μ_{rt}, $r,t=1,2$.
iii) How to choose the receive filters $\boldsymbol{Q}_t$, $t=1,2$.
iv) How to compute ζ_{rt}, $r,t=1,2$.

Having completed these steps, the procedure of computing the filters $\boldsymbol{\Psi}^{\dagger}_{rt}, r,t=1,2$, is similar for all cases. Later, we will show that this scheme decomposes the system into two noninterfering multiple-access channels.

Scheme II–Case I: $M_1 \geq M_2 \geq N_1 \geq N_2$:

In this case, the system is irreducible Type II. Therefore, the equivalent system is the same as the original system, i.e., $N'_r = N_r$, $r=1,2$ and $M'_t = M_t$, $t=1,2$. Using the above parameters, we choose μ'_{rt}, $r,t=1,2$, subject to (80)–(85). Similar to Scheme I–Case I, we have $\mu_{rt}=\mu'_{rt}, r,t=1,2$. $\boldsymbol{Q}_1$ and $\boldsymbol{Q}_2$ are randomly chosen from $\mathcal{OC}^{N_1\times(\mu_{11}+\mu_{12})}$ and $\mathcal{OC}^{N_2\times(\mu_{21}+\mu_{22})}$, respectively.

According to the definition of $\zeta_{rt}, r,t=1,2$, it is easy to see that

$$\begin{aligned}\zeta_{11}&=\mu_{21}+\mu_{22}, &\zeta_{12}&=\mu_{21}+\mu_{22},\\ \zeta_{21}&=\mu_{12}+\mu_{11}, &\zeta_{22}&=\mu_{11}+\mu_{12}.\end{aligned}\tag{86}$$

Scheme II–Case II: $M_1 \geq N_1 > M_2 \geq N_2$:

In this case, at the receiver one, the signal coming from transmitter two does not have any component in the (N_1-M_2)-dimensional subspace $\mathrm{N}(\boldsymbol{H}^{\dagger}_{12})$. This subspace can be exploited to increase the number of data streams sent from transmitter one to receiver one by N_1-M_2 without restricting the available signaling space at transmitter two and at receiver two. Consequently, the system is reduced to a system with $(M'_1,M'_2,N'_1,N'_2)=(M_1-\{N_1-M_2\},M_2,N_1-\{N_1-M_2\},N_2)$ antennas, or

$$M'_1=M_1+M_2-N_1,\ M'_2=M_2,\ N'_1=M_2,\ N'_2=N_2.\tag{87}$$

It is easy to see that $M'_1 \geq M'_2 \geq N'_1 \geq N'_2$, i.e., the original system is reducible to Type II. We choose μ'_{rt}, $r,t=1,2$, subject to (80)–(85). Then, we have

$$\mu_{11}=\mu'_{11}+N_1-M_2,\ \mu_{12}=\mu'_{12},\ \mu_{21}=\mu'_{21},\ \mu_{22}=\mu'_{22}.\tag{88}$$

$\boldsymbol{Q}_1$ is chosen as

$$\boldsymbol{Q}_1\in\mathcal{OC}^{N_1\times(\mu_{11}+\mu_{21})},\quad \boldsymbol{Q}_1=[\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2]\tag{89}$$

where

$$\boldsymbol{\Sigma}_1\in\mathcal{OC}^{N_1\times(N_1-M_2)},\quad \boldsymbol{\Sigma}_1\in\mathrm{N}(\boldsymbol{H}^{\dagger}_{12})\tag{90}$$

$$\boldsymbol{\Sigma}_2\in\mathcal{OC}^{N_1\times(\mu'_{11}+\mu_{12})},\quad \boldsymbol{\Sigma}_2\perp\boldsymbol{\Sigma}_1.\tag{91}$$

$\boldsymbol{Q}_2$ is randomly selected from $\mathcal{OC}^{N_2\times(\mu_{21}+\mu_{22})}$.

It is easy to see that

$$\zeta_{11}=\mu_{21}+\mu_{22},\ \zeta_{12}=\mu_{21}+\mu_{22},\ \zeta_{21}=\mu_{12}+\mu_{11},\ \zeta_{22}=\mu'_{11}+\mu_{12}.\tag{92}$$

Scheme II–Case III: $M_1 \geq N_1 \geq N_2 > M_2$ *and* $M_1+M_2 \geq N_1+N_2$:

In this case, at receiver one, the signal coming from transmitter two has no component in the (N_1-M_2)-dimensional subspace $\mathrm{N}(\boldsymbol{H}^{\dagger}_{12})$. This subspace can be exploited to increase the number of data streams sent from transmitter one to receiver one by N_1-M_2 without restricting the available signaling space at transmitter two and at receiver two. In addition, at receiver two, the signal coming from transmitter two has no component in the (N_2-M_2)-dimensional subspace $\mathrm{N}(\boldsymbol{H}^{\dagger}_{22})$. This subspace can be exploited to increase the number of data streams sent from transmitter one to receiver two by N_2-M_2, without restricting the available signaling space at transmitter two and at receiver one. Therefore, the reduced system has (M'_1,M'_2,N'_1,N'_2) antennas, where

$$M'_1=M_1+2M_2-N_1-N_2,\ M'_2=M_2,\ N'_1=M_2,\ N'_2=M_2.\tag{93}$$

$M'_1 \geq M'_2 \geq N'_1 \geq N'_2$. Therefore, the original system is reducible to Type II. After choosing μ'_{rt}, $r,t=1,2$, subject to (80)–(85), we have

$$\mu_{11}=\mu'_{11}+N_1-M_2,\ \mu_{12}=\mu'_{12},\ \mu_{21}=\mu'_{21}+N_2-M_2,\ \mu_{22}=\mu'_{22}.\tag{94}$$

$\boldsymbol{Q}_1$ is chosen as

$$\boldsymbol{Q}_1\in\mathcal{OC}^{N_1\times(\mu_{11}+\mu_{21})},\quad \boldsymbol{Q}_1=[\boldsymbol{\Sigma}_1,\boldsymbol{\Sigma}_2]\tag{95}$$

where

$$\boldsymbol{\Sigma}_1\in\mathcal{OC}^{N_1\times(N_1-M_2)}\quad \boldsymbol{\Sigma}_1\in\mathrm{N}(\boldsymbol{H}^{\dagger}_{21})\tag{96}$$

$$\boldsymbol{\Sigma}_2\in\mathcal{OC}^{N_1\times(\mu'_{11}+\mu_{12})}\quad \boldsymbol{\Sigma}_2\perp\boldsymbol{\Sigma}_1.\tag{97}$$

$\boldsymbol{Q}_2$ is chosen as

$$\boldsymbol{Q}_2\in\mathcal{OC}^{N_2\times(\mu_{21}+\mu_{22})},\quad \boldsymbol{Q}_2=[\boldsymbol{\Sigma}_3,\boldsymbol{\Sigma}_4]\tag{98}$$

where

$$\boldsymbol{\Sigma}_3\in\mathcal{OC}^{N_2\times(N_2-M_2)},\quad \boldsymbol{\Sigma}_3\in\mathrm{N}(\boldsymbol{H}^{\dagger}_{22})\tag{99}$$

$$\boldsymbol{\Sigma}_4\in\mathcal{OC}^{N_2\times(\mu'_{21}+\mu_{22})},\quad \boldsymbol{\Sigma}_4\perp\boldsymbol{\Sigma}_3.\tag{100}$$

Therefore, we have

$$\zeta_{11}=\mu_{21}+\mu_{22},\ \zeta_{12}=\mu'_{21}+\mu_{22},\ \zeta_{21}=\mu_{12}+\mu_{11},\ \zeta_{22}=\mu'_{11}+\mu_{12}.\tag{101}$$

Scheme II–Case IV: $N_1 > M_1 \geq M_2 \geq N_2$ *and* $M_1+M_2 \geq N_1+N_2$:

In this case, i) (N_1-M_2)-dimensional subspace $\mathrm{N}(\boldsymbol{H}^{\dagger}_{12})$ is utilized to increase the number of data streams sent from transmitter one to receiver one by (N_1-M_2), ii) (N_1-M_1)-dimensional subspace $\mathrm{N}(\boldsymbol{H}^{\dagger}_{11})$ is utilized to increase the number of data streams sent from transmitter two to receiver one by $(N_1 - M_1)$. Therefore, we have

$$\begin{aligned} M_1' &= M_1 + M_2 - N_1, \quad M_2' = M_1 + M_2 - N_1, \\ N_1' &= M_1 + M_2 - N_1, \quad N_2' = N_2, \\ \mu_{11} &= \mu_{11}' + N_1 - M_2, \quad \mu_{12} = \mu_{12}' + N_1 - M_1, \\ \mu_{21} &= \mu_{21}', \quad \mu_{22} = \mu_{22}' \end{aligned} \tag{102}$$

where $M_1' \geq M_2' \geq N_1' \geq N_2'$, i.e., the original system is reducible to Type II. $\boldsymbol{Q}_1$ is chosen as

$$\boldsymbol{Q}_1 \in \mathcal{OC}^{N_1 \times (\mu_{11}+\mu_{21})}, \quad \boldsymbol{Q}_1 = [\boldsymbol{\Sigma}_1, \boldsymbol{\Sigma}_2] \tag{103}$$

where

$$\boldsymbol{\Sigma}_1 \in \mathcal{OC}^{N_1 \times (N_1 - M_2 + N_1 - M_2)}, \quad \boldsymbol{\Sigma}_1 \in \mathrm{N}(\boldsymbol{H}_{12}^\dagger) \cup \mathrm{N}(\boldsymbol{H}_{11}^\dagger) \tag{104}$$

$$\boldsymbol{\Sigma}_2 \in \mathcal{OC}^{N_1 \times (\mu_{11}'+\mu_{12}')}, \quad \boldsymbol{\Sigma}_2 \perp \boldsymbol{\Sigma}_1. \tag{105}$$

$\boldsymbol{Q}_2$ is randomly selected from $\mathcal{OC}^{N_2 \times (\mu_{21}+\mu_{22})}$. In this case, we have

$$\zeta_{11} = \mu_{21} + \mu_{22}, \ \zeta_{12} = \mu_{21} + \mu_{22}, \ \zeta_{21} = \mu_{12}' + \mu_{11}, \ \zeta_{22} = \mu_{11}' + \mu_{12}. \tag{106}$$

The next steps of the algorithm are the same for all above cases. We define

$$\widetilde{\boldsymbol{H}}_{rt} = \boldsymbol{Q}_r^\dagger \boldsymbol{H}_{rt}, \quad r,t = 1,2. \tag{107}$$

$\boldsymbol{\Psi}_{rt} \in \mathcal{OC}^{M_r \times (M_r - \zeta_{rt})}$, $r,t = 1,2$, are chosen such that

$$\boldsymbol{\Psi}_{11} \perp \widetilde{\boldsymbol{H}}_{21}^\dagger \tag{108}$$

$$\boldsymbol{\Psi}_{21} \perp \widetilde{\boldsymbol{H}}_{11}^\dagger \tag{109}$$

$$\boldsymbol{\Psi}_{12} \perp \widetilde{\boldsymbol{H}}_{22}^\dagger \tag{110}$$

$$\boldsymbol{\Psi}_{22} \perp \widetilde{\boldsymbol{H}}_{12}^\dagger. \tag{111}$$

According to the definition of ζ_{rt}, we can always choose such matrices. Clearly, any signal passed through the filters $\boldsymbol{\Psi}_{11}^\dagger$ and $\boldsymbol{\Psi}_{12}^\dagger$ has no interference at the output of the filter $\boldsymbol{Q}_2$. Similarly, any signal passed through the filters $\boldsymbol{\Psi}_{21}^\dagger$ and $\boldsymbol{\Psi}_{22}^\dagger$ has no interference at the output of the filter $\boldsymbol{Q}_2$. We define

$$\overline{\boldsymbol{H}}_{rt} = \widetilde{\boldsymbol{H}}_{rt} \boldsymbol{\Psi}_{rt}, \quad r,t = 1,2 \tag{112}$$

and

$$\tilde{\boldsymbol{w}}_r = \boldsymbol{Q}_r^\dagger \boldsymbol{w}_r, \quad r,t = 1,2. \tag{113}$$

This system is decomposed into two noninterfering multiple-access channels: i) the multiple-access channel viewed by receiver one with channels $\overline{\boldsymbol{H}}_{11}$ and $\overline{\boldsymbol{H}}_{12}$, modeled by (see Fig. 4)

$$\tilde{\boldsymbol{y}}_1 = \overline{\boldsymbol{H}}_{11} \boldsymbol{s}_{11} + \overline{\boldsymbol{H}}_{12} \boldsymbol{s}_{12} + \tilde{\boldsymbol{w}}_1 \tag{114}$$

and ii) the multiple-access channel viewed by receiver two with channels $\overline{\boldsymbol{H}}_{21}$ and $\overline{\boldsymbol{H}}_{22}$, modeled by (see Fig. 4)

$$\tilde{\boldsymbol{y}}_2 = \overline{\boldsymbol{H}}_{21} \boldsymbol{s}_{21} + \overline{\boldsymbol{H}}_{22} \boldsymbol{s}_{12} + \tilde{\boldsymbol{w}}_2. \tag{115}$$

Acknowledgment

The authors would like to thank the anonymous reviewers for thoughtful comments.

References

[1] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–335, 1998.
[2] İ. E. Telatar, "Capacity of multiple-antenna Gaussian channels," *Europ. Trans. Telecommun.*, pp. 585–595, Nov. 1999.
[3] W. Yu and W. Rhee, "Degrees of freedom in multiple-user spatial multiplex systems with multiple antennas," *IEEE Trans. Commun.*, vol. 54, no. 10, pp. 1747–1753, Oct. 2006.
[4] S. A. Jafar, "Degrees of freedom in distributed mimo communications," in *Proc. IEEE Communication Theory Workshop*, UT, 2005.
[5] S. Vishwanath and S. A. Jafar, "On the capacity of vector Gaussian interference channels," in *Proc. IEEE Information Theory Workshop*, Austin, TX, 2004, pp. 365–369.
[6] M. Costa and A. El Gamal, "The capacity region of the discrete memoryless interference channel with strong interference," *IEEE Trans. Inf. Theory*, vol. 33, no. 5, pp. 710–711, Sep. 1987.
[7] A. Carleial, "A case where interference does not reduce capacity (corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-21, no. 5, pp. 569–570, Sep. 1975.
[8] X. Shang, B. Chen, and M. J. Gans, "On the achievable sum rate for MIMO interference channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 4313–4320, Sep. 2006.
[9] 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.
[10] S. Shamai (Shitz) and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Proc. IEEE Vehicular Technology Conf.*, May 2001, vol. 3, pp. 1745–1749.
[11] G. J. Foschini, H. Huang, K. Karakayali, R. A. Valenzuela, and S. Venkatesan, "The value of coherent base station coordination," in *Proc. Conf. Information Sciences and Systems (CISS)*, Baltimore, MD, Mar. 2005.
[12] A. Høst-Madsen, "Capacity bounds for cooperative diversity," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1522–1544, Apr. 2006.
[13] M. A. Maddah-Ali, S. A. Motahari, and A. K. Khandani, "Signaling over MIMO multiple-base systems: Combination of multiple-access and broadcast schemes," in *Proc. IEEE Int. Symp. Information Theory*, Seattle, WA, Jul. 2006, pp. 2104–2108.
[14] M. A. Maddah-Ali, S. A. Motahari, and A. K. Khandani, Communication Over X Channel: Signalling and Multiplexing Gain Univ. Waterloo, Waterloo, ON, Canada, Tech. Rep. UW-ECE-2006–12, 2006.
[15] S. A. Jafar, Degrees of Freedom On the MIMO X Channel—The Optimality of the MMK Scheme, 2006, [Online]. Available: http://arxiv.org/abs/cs.IT/0607099
[16] N. Devroye and M. Sharif, "The multiplexing gain of MIMO X-channels with partial transmit side-information," in *Proc. IEEE Int. Symp. Information Theory*, Nice, France, Jun. 2007, pp. 111–115.
[17] S. A. Jafar and S. Shamai (Shitz), "Degrees of freedom region for the MIMO X channel," *IEEE Trans. Inf. Theory*, 2007, submitted for publication.
[18] S. Shamai (Shitz) and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1302–1327, May 2001.
[19] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR power offset in multiantenna communication," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.
[20] N. Jindal, "High SNR analysis of MIMO broadcast channels," in *Proc. IEEE Int. Symp. Information Theory*, Adelaide, Australia, Sep. 2005, pp. 2310–2314.
[21] M. A. Maddah-Ali, M. Ansari, and A. K. Khandani, "An efficient signaling scheme for MIMO broadcast systems: Design and performance evaluation," *IEEE Trans. Inf. Theory*, submitted for publication.
[22] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 3, pp. 439–441, May 1983.
[23] R. G. Horn and C. A. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1985.
[24] R. G. Gallager, *Information Theory and Reliable Communication*. New York: Wiley, 1968.
[25] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.
[26] P. Viswanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.
[27] W. Yu and J. Cioffi, "Sum capacity of vector Gaussian broadcast channels," *IEEE Trans. Inf. Theory*, submitted for publication.
[28] W. Yu, "A dual decomposition approach to the sum power Gaussian vector multiple access channel sum capacity problem," in *Proc. 37th Annu. Conf. Information Sciences and Systems (CISS)*, Baltimore, MD, Mar. 2003.