MIMO Broadcast Channel Rate Region with Linear Precoding at High SNR without Full Multiplexing

Paul de Kerret*, Michael Joham†, Wolfgang Utschick†, and Rudolf Mathar*
* Institute for Theoretical Information Technology, RWTH Aachen University,
52056 Aachen, Germany, email:{dekerret,mathar}@ti.rwth-aachen.de
† Associate Institute for Signal Processing, Technische Universität München,
80290 Munich, Germany, email:{joham,utschick}@tum.de

***Abstract*—In this paper, the rate region of the two user MIMO broadcast channel (BC) with linear filtering at high signal-to-noise ratio (SNR) is studied when time sharing is not available and the transmitter has fewer antennas than the sum of the receiving antennas. To reach the boundary of the rate region, the sum rate is maximized subject to a rate ratio constraint. Furthermore, the sum rate is approximated as an affine function of the logarithm of the SNR and the two parameters of this approximation, which are the multiplexing gain (MG) and the rate offset (RO), are derived. This leads directly to the asymptotic rate region, particularly interesting because it is obtained in simple analytical form and offers a good approximation at high but finite SNR. We then consider the rate region boundary at finite SNR and derive algorithmic bounds for it, which are accurate even at intermediate SNR.**

I. Introduction

We consider a two user broadcast channel (BC) where both users have several antennas and the base station (BS) is assumed to have fewer antennas than the sum of the receiving antennas. The transmitter and the receivers have perfect channel state information and the power available at the transmitter is very large. Moreover, time sharing is not available.

At arbitrary SNR, the capacity region is then known to be achievable with dirty paper coding (DPC) [1], [2] and globally optimum algorithms are available to maximize the weighted sum rate [3]. Still, DPC requires a very demanding implementation [4], while linear precoding is a suboptimal alternative with good performance and low complexity. Thus, we assume in the following that the BS applies linear precoding, in which case only lower bounds for the convex hull have been obtained algorithmically [5], [6].

At high-SNR, a common approximation is to let the SNR tend to infinity and to write the sum rate R as

$$R = \mathrm{M_G}\left(\log(\mathrm{SNR}) - \mathrm{R_O}\right),$$

where $\mathrm{M_G}$ and $\mathrm{R_O}$ are the multiplexing gain (MG) and the rate offset (RO), respectively.

With more transmitting antennas than the sum of the receiving antennas, analytical expressions have been derived for the optimal weighted sum rate of linear precoding [7], [8]. The maximization of the sum rate subject to a *rate ratio* constraint is also studied in [9], [10], and leads to the high-SNR rate region when time sharing is not available.

However, no result exists at high-SNR when the transmitter has fewer antennas than the sum of the receiving antennas, and the aim of this work is to fill this gap.

The main contributions read as follows. First, the asymptotically optimal stream allocation, which leads to the MG and the RO, is derived at every point of the boundary of the rate region. Second, it is shown that the precoding matrices derived are very close to optimal and can be used to derive inner and outer bounds for the rate region boundary.

In the sequel, the calculations are based on the rate duality between the MIMO BC and a dual MIMO Multiple Access Channel (MAC) with the same sum power constraint [11], which allows us to study the rate region in the dual MAC.

In Section II, we introduce the system model and the optimization problem. The MG is then studied in Section III and the RO in Section IV. Finally, the results are applied to describe the rate region at finite SNR in Section V, and graphical illustrations are given in Section VI.

Notation: The operators $\|\cdot\|_{\mathrm{F}}$, $|\cdot|$, $(\cdot)^{\mathrm{H}}$, $\log(\cdot)$, and $\lceil\cdot\rceil$ denote the Frobenius norm, the determinant operator, the Hermitian transposition, the logarithm base 2 and the ceiling operator, respectively. We also write *streams* instead of *independent data streams* and w.l.o.g. for *without loss of generality*.

II. System Model

A. Rate Expressions

We consider a BC with two users, denoted as user 1 and user 2, having r_1 and r_2 antennas, respectively, while the BS is equipped with t antennas and $t < r_1 + r_2$. We denote the antenna configuration of the users by $\boldsymbol{r} = (r_1, r_2)$ and the stream allocation by $\boldsymbol{b} = (b_1, b_2)$, where b_i is the number of streams allocated to user i. The stream allocation always verifies $b_1 + b_2 \leq t$, $b_1 \leq r_1$, and $b_2 \leq r_2$. The power available at the BS is given by P and is normalized to the variance of the noise, such that it is assimilated to the SNR. Using the duality between the rate region of the BC channel and the dual MAC [11], we consider the transmission in the dual MAC in which the two users transmit to the BS with the sum power constraint of the original BC channel. It means that P has to be split into P_1 and P_2, which correspond to the power allocated to user 1 and 2, respectively. The channel seen by user i is given by $\mathbf{H}_i \in \mathbb{C}^{t\times r_i}$ and is assumed to be full rank and perfectly known at the BS and at both users. Each element of the

978–1–4244–6014–4/10/$26.00

channel is generated randomly from an independent identically distributed standard Gaussian distribution, and the same holds for every element of the noise vector at the receiver. When user i applies the full rank precoding matrix $\mathbf{T}_i \in \mathbb{C}^{r_i \times b_i}$, the rates of user 1 and 2 are given by [11]

$$R_1 = \log\left|\mathbf{I}+\mathbf{T}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}(\mathbf{I}-\mathbf{H}_2\mathbf{T}_2(\mathbf{I}+\mathbf{T}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2\mathbf{T}_2)^{-1}\mathbf{T}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}})\mathbf{H}_1\mathbf{T}_1\right|,$$
$$R_2 = \log\left|\mathbf{I}+\mathbf{T}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}(\mathbf{I}-\mathbf{H}_1\mathbf{T}_1(\mathbf{I}+\mathbf{T}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1\mathbf{T}_1)^{-1}\mathbf{T}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}})\mathbf{H}_2\mathbf{T}_2\right|. \tag{1}$$

We now decompose the precoding matrices $\mathbf{T}_i = \sqrt{P_i/b_i}\bar{\mathbf{T}}_i$, where $\bar{\mathbf{T}}_i$ is the *normalized precoding matrix* (NP matrix), such that $\|\bar{\mathbf{T}}_i\|_{\mathrm{F}}^2 = b_i$. Since we consider the high-SNR regime, every non-zero eigenvalue of the transmit covariance matrix can be assumed to be very large. Thus, the identity in the inverse term in (1) can be neglected, and we get

$$R_1' = \log\left|\mathbf{I}+\mathbf{T}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}(\mathbf{I}-\mathbf{H}_2\bar{\mathbf{T}}_2(\bar{\mathbf{T}}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2\bar{\mathbf{T}}_2)^{-1}\bar{\mathbf{T}}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}})\mathbf{H}_1\mathbf{T}_1\right|,$$
$$R_2' = \log\left|\mathbf{I}+\mathbf{T}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}(\mathbf{I}-\mathbf{H}_1\bar{\mathbf{T}}_1(\bar{\mathbf{T}}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1\bar{\mathbf{T}}_1)^{-1}\bar{\mathbf{T}}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}})\mathbf{H}_2\mathbf{T}_2\right|.$$

The high-SNR assumption is now further used to neglect the offset identities inside the determinants. With this approximation, the high-SNR approximated rates read as

$$R_1'' = \log\left|\mathbf{T}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}(\mathbf{I}-\mathbf{H}_2\bar{\mathbf{T}}_2(\bar{\mathbf{T}}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2\bar{\mathbf{T}}_2)^{-1}\bar{\mathbf{T}}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}})\mathbf{H}_1\mathbf{T}_1\right|,$$
$$R_2'' = \log\left|\mathbf{T}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}(\mathbf{I}-\mathbf{H}_1\bar{\mathbf{T}}_1(\bar{\mathbf{T}}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1\bar{\mathbf{T}}_1)^{-1}\bar{\mathbf{T}}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}})\mathbf{H}_2\mathbf{T}_2\right|. \tag{2}$$

In the sequel, we consider only the rates R_1'' and R_2'' from (2) instead of the exact ones from (1). The error due to the high-SNR approximation is discussed in [12], [13], and it is shown that the approximation is already very accurate at $P = 30$ dB.

We now define *the projected channels* as

$$\bar{\mathbf{H}}_1(\bar{\mathbf{T}}_2) := (\mathbf{I}_t - \mathbf{H}_2\bar{\mathbf{T}}_2(\bar{\mathbf{T}}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2\bar{\mathbf{T}}_2)^{-1}\bar{\mathbf{T}}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}})\mathbf{H}_1,$$
$$\bar{\mathbf{H}}_2(\bar{\mathbf{T}}_1) := (\mathbf{I}_t - \mathbf{H}_1\bar{\mathbf{T}}_1(\bar{\mathbf{T}}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1\bar{\mathbf{T}}_1)^{-1}\bar{\mathbf{T}}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}})\mathbf{H}_2,$$

and the *rate shifts* as

$$\begin{aligned} c_1(\bar{\mathbf{T}}_1, \bar{\mathbf{T}}_2) &:= \left|\bar{\mathbf{T}}_1^{\mathrm{H}}\bar{\mathbf{H}}_1^{\mathrm{H}}(\bar{\mathbf{T}}_2)\bar{\mathbf{H}}_1(\bar{\mathbf{T}}_2)\bar{\mathbf{T}}_1\right|^{-\frac{1}{b_1}}, \\ c_2(\bar{\mathbf{T}}_1, \bar{\mathbf{T}}_2) &:= \left|\bar{\mathbf{T}}_2^{\mathrm{H}}\bar{\mathbf{H}}_2^{\mathrm{H}}(\bar{\mathbf{T}}_1)\bar{\mathbf{H}}_2(\bar{\mathbf{T}}_1)\bar{\mathbf{T}}_2\right|^{-\frac{1}{b_2}}. \end{aligned} \tag{3}$$

We can also rewrite the rate shift as

$$c_1(\bar{\mathbf{T}}_1, \bar{\mathbf{T}}_2) = \left|\bar{\mathbf{T}}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1\bar{\mathbf{T}}_1\right|^{-\frac{1}{b_1}} \frac{\left|\bar{\mathbf{T}}_2^{\mathrm{H}}\bar{\mathbf{H}}_2^{\mathrm{H}}(\bar{\mathbf{T}}_1)\bar{\mathbf{H}}_2(\bar{\mathbf{T}}_1)\bar{\mathbf{T}}_2\right|^{-\frac{1}{b_1}}}{\left|\bar{\mathbf{T}}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2\bar{\mathbf{T}}_2\right|^{-\frac{1}{b_1}}},$$
$$c_2(\bar{\mathbf{T}}_1, \bar{\mathbf{T}}_2) = \left|\bar{\mathbf{T}}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2\bar{\mathbf{T}}_2\right|^{-\frac{1}{b_2}} \frac{\left|\bar{\mathbf{T}}_1^{\mathrm{H}}\bar{\mathbf{H}}_1^{\mathrm{H}}(\bar{\mathbf{T}}_2)\bar{\mathbf{H}}_1(\bar{\mathbf{T}}_2)\bar{\mathbf{T}}_1\right|^{-\frac{1}{b_2}}}{\left|\bar{\mathbf{T}}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1\bar{\mathbf{T}}_1\right|^{-\frac{1}{b_2}}}. \tag{4}$$

The approximated rates from (2) are then given by

$$\begin{aligned} R_1'' &= b_1\left(\log(P_1) - \log(b_1) - \log(c_1(\bar{\mathbf{T}}_1, \bar{\mathbf{T}}_2))\right), \\ R_2'' &= b_2\left(\log(P_2) - \log(b_2) - \log(c_2(\bar{\mathbf{T}}_1, \bar{\mathbf{T}}_2))\right). \end{aligned} \tag{5}$$

B. Optimization Problem

We consider the maximization of the sum rate subject to a *rate ratio constraint*, i.e., to maximize the sum rates $R'' := R_1'' + R_2''$ subject to a given ratio between the rates R_1'' and R_2''. This constraint is expressed by means of the nonnegative *rate coefficients* $\boldsymbol{\gamma} := (\gamma_1, \gamma_2)$ such that $R_1''/\gamma_1 = R_2''/\gamma_2$. The rates coefficients are normalized as $\gamma_1 + \gamma_2 = 1$, which implies $R_1'' = \gamma_1 R''$ and $R_2'' = \gamma_2 R''$. Note that the knowledge of γ_1, γ_2 or $\boldsymbol{\gamma}$ is equivalent. Similarly, once the rate ratio constraint is fulfilled, the knowledge of R_1'', R_2'' or R'' is also equivalent.

Using these definitions, the optimization problem is

$$\max_{\bar{\mathbf{T}}_1, \bar{\mathbf{T}}_2, P_1, P_2, b_1, b_2} R'' \quad \text{subject to: } \frac{R_1''}{\gamma_1} = \frac{R_2''}{\gamma_2},\ P_1 + P_2 = P,$$
$$P_i \geq 0, \quad \|\bar{\mathbf{T}}_i\|_{\mathrm{F}}^2 = b_i, \quad i = 1, 2. \tag{6}$$

In this work, we will mainly consider that P tends to infinity and use the following approximation for the sum rate as a function of the available power:

$$R''(\boldsymbol{\gamma}) \approx R_\infty(\boldsymbol{\gamma}) = \mathrm{M_G}(\boldsymbol{\gamma})\left(\log(P) - \mathrm{R_O}(\boldsymbol{\gamma})\right), \tag{7}$$

where $\mathrm{M_G}(\boldsymbol{\gamma})$ and $\mathrm{R_O}(\boldsymbol{\gamma})$ are the multiplexing gain (MG) and the rate offset (RO), respectively. We start by computing the maximal MG and focus then on the RO.

III. Multiplexing Gain Study

To achieve the optimal MG, the stream allocation has to be optimized. We start by recalling some results from [10] for a given stream allocation $\boldsymbol{b}$ and then use these results to obtain the MG region with optimal stream allocation.

A. Multiplexing Gain for a Fixed Stream Allocation $\boldsymbol{b}$.

Theorem 1. *[10] For given rate coefficients $\boldsymbol{\gamma}$ and stream allocation $\boldsymbol{b}$, only one user has a power allocation scaling linearly in P. He is called the* limiting user *and denoted as user ℓ, while the other one, called the* non-limiting *user and denoted as user $n\ell$, has a sub-linear power allocation scaling with P raised to the exponent $b_\ell\gamma_{n\ell}/(b_{n\ell}\gamma_\ell) < 1$. The limiting user is the user with the largest quotient γ_i/b_i and is hence user 2 if $\gamma_1 < b_1/(b_1+b_2)$, and user 1 if $\gamma_1 > b_1/(b_1+b_2)$. The rate coefficients corresponding to the equality are called the* transition coefficients of the stream allocation $\boldsymbol{b}$ *and are denoted by $\boldsymbol{\gamma}_{\mathrm{tr}}(\boldsymbol{b}) = (\gamma_{1,\mathrm{tr}}(\boldsymbol{b}), \gamma_{2,\mathrm{tr}}(\boldsymbol{b}))$. The maximal MG of the stream allocation $\boldsymbol{b}$ is achieved only at $\boldsymbol{\gamma}_{\mathrm{tr}}(\boldsymbol{b})$. The MG for arbitrary $\boldsymbol{\gamma}$ reads as follows.*

$$\begin{aligned} &\text{If } \gamma_1 < \gamma_{1,\mathrm{tr}}(\boldsymbol{b}) = \frac{b_1}{b_1+b_2}: \quad M_G(\boldsymbol{\gamma}) = \frac{b_2}{\gamma_2}, \\ &\text{If } \gamma_1 = \gamma_{1,\mathrm{tr}}(\boldsymbol{b}) = \frac{b_1}{b_1+b_2}: \quad M_G(\boldsymbol{\gamma}) = b_1 + b_2, \\ &\text{If } \gamma_1 > \gamma_{1,\mathrm{tr}}(\boldsymbol{b}) = \frac{b_1}{b_1+b_2}: \quad M_G(\boldsymbol{\gamma}) = \frac{b_1}{\gamma_1}. \end{aligned} \tag{8}$$

Proof: The MG does not depend on the rate shifts, which we thus denote for the proof with the simplified notation c_1

and c_2. We now rewrite the rate ratio constraint using (5) as

$$\frac{R_1''}{\gamma_1} = \frac{R_2''}{\gamma_2} \quad \Leftrightarrow \left(\frac{P_1}{b_1 c_1}\right)^{\frac{b_1}{\gamma_1}} = \left(\frac{P_2}{b_2 c_2}\right)^{\frac{b_2}{\gamma_2}}. \tag{9}$$

Inserting (9) in the power constraint gives

$$P_1 + P_2 = P_1 + b_2 c_2 P_1^{\frac{b_1\gamma_2}{\gamma_1 b_2}} (b_1 c_1)^{-\frac{b_1\gamma_2}{\gamma_1 b_2}} = P \tag{10}$$

Considering w.l.o.g. that $\gamma_1 > \gamma_{1,\text{tr}}(\boldsymbol{b})$, it implies that $\frac{b_1\gamma_2}{\gamma_1 b_2} < 1$. Letting P tend to infinity in (10), P_1 tends also to infinity such that the term with P_1 raised to the largest exponent is dominant and P_1 scales linearly with P. The scaling of P_2 is then derived from (9) and the MG follows directly. ■

The MG region is easily seen to be a rectangle of dimension $b_1 \times b_2$. The vertex corresponds to the transition coefficients $\boldsymbol{\gamma}_{\text{tr}}(\boldsymbol{b})$ and is the only point at which both users have a power allocation scaling linearly in P. If the ray starting from the origin associated with a given $\boldsymbol{\gamma}$ intersects the rectangle on the vertical part, user 1 is the limiting user (and has hence a power allocation scaling linearly in P), and if it is on the horizontal part, user 2 is the limiting user.

B. Multiplexing Gain for Particular Antenna Configurations

To derive results for the general antenna configuration, we start by studying two particular antenna configurations: $t \geq r_1 + r_2$ and $t \leq \min(r_1, r_2)$. The configuration where $t \geq r_1 + r_2$ has been studied in [10] and one of the theorems is recalled here.

Theorem 2. *[10] Let $t \geq r_1 + r_2$. The maximal $M_G(\boldsymbol{\gamma})$ and the stream allocation $\boldsymbol{b}$ to achieve it read as follows.*

For $\gamma_1 \leq \gamma_{1,\text{tr}}(\boldsymbol{r})$, $M_G(\boldsymbol{\gamma}) = \frac{r_2}{\gamma_2}$, if: $b_1 \geq \left\lceil \frac{\gamma_1 r_2}{\gamma_2} \right\rceil, b_2 = r_2$.

For $\gamma_1 \geq \gamma_{1,\text{tr}}(\boldsymbol{r})$, $M_G(\boldsymbol{\gamma}) = \frac{r_1}{\gamma_1}$, if: $b_2 \geq \left\lceil \frac{\gamma_2 r_1}{\gamma_1} \right\rceil, b_1 = r_1$.

We now consider the particular case $t \leq \min(r_1, r_2)$.

Theorem 3. *Let $t \leq \min(r_1, r_2)$. The maximal MG is t and is achieved only at the rate coefficients $\gamma_{1,\max}(i) := i/t, i \in \{0, \ldots, t\}$. The minimal MG is $t-1$ and is achieved only at the rate coefficients $\gamma_{1,\min}(i) := (i-1)/(t-1), i \in \{2, \ldots, t-2\}$ with the stream allocation $\boldsymbol{b} = (i-1, t-i)$. For all $i \in \{1, \ldots, t-1\}$, the MG reads otherwise as follows.*

$$\begin{aligned} &\textit{If } \gamma_1 \in]\gamma_{1,\min}(i), \gamma_{1,\max}(i)], && M_G(\boldsymbol{\gamma}) = b_2/\gamma_2. \\ &\textit{If } \gamma_1 \in [\gamma_{1,\max}(i), \gamma_{1,\min}(i+1)[, && M_G(\boldsymbol{\gamma}) = b_1/\gamma_1. \end{aligned} \tag{11}$$

The asymptotically optimal stream allocation is then $\boldsymbol{b} = (i, t-i)$ for $\gamma_1 \in]\gamma_{1,\min}(i), \gamma_{1,\min}(i+1)[, i \in \{1, \ldots, t-1\}$.

Proof: Since only the limiting user has a power allocation scaling linearly in P, see Theorem 1, the MG is given by

$$\mathrm{M_G}(\boldsymbol{\gamma}) = \min_{b_1}\left(\frac{b_1}{\gamma_1}, \frac{b_2}{\gamma_2}\right), \text{with } b_1 \in \{0, \ldots, t\}, b_2 = t - b_1.$$

If we relax the constraint of b_1 being an integer and let it be a real number, the optimal solution is obtained when the two terms in the $\min(\cdot)$ are equal and the solution is $b_1 = \gamma_1 t$ (and $b_2 = \gamma_2 t$) with a MG achieved always equal to t.

The rate coefficients corresponding to integer values are the rate coefficients $\gamma_{1,\max}(i) = \frac{i}{t}, \forall i \in \{0, \ldots, t\}$, at which the number of streams transmitted by user 1 is i and the maximum MG is achieved. For the other values of γ_1, we need to choose the integer value for the number of transmitted streams which leads to the largest MG.

Let $\gamma_1 \in]\gamma_{1,\max}(i-1), \gamma_{1,\max}(i)[, i \in \{1, \ldots, t\}$. The number of streams transmitted by user 1 has to be chosen between the number of streams transmitted by user 1 at $\gamma_{1,\max}(i-1)$ and at $\gamma_{1,\max}(i)$, i.e., $i-1$ and i, respectively, while user 2 always transmits $t - b_1$ streams. When user 1 transmits $i-1$ streams, the MG is $(i-1)/\gamma_1$, because user 1 is then the limiting user. When user 1 transmits i streams, the MG is $(t-i)/(1-\gamma_1)$ because user 2 is then the limiting user. There is a unique intersection, since the first expression decreases monotonically in γ_1 and the second one increases monotonically in γ_1, which occurs at $\gamma_{1,\min}(i)$. The MG achieved there is the smallest and equal to $t-1$. This holds for every $i \in \{1, \ldots, t\}$ and concludes the proof. ■

C. Multiplexing Gain for Arbitrary Antenna Configurations

Theorem 4. *Let t, r_1, and r_2 be arbitrarily given.*

If $\gamma_1 < \phi_1 := r_1/t$ or $\gamma_1 > \phi_2 := (t-r_2)/t$, the MG and $\boldsymbol{b}$ are then the same as if $t \geq r_1 + r_2$ and are given by Theorem 2. If $\gamma_1 \in [\phi_1, \phi_2]$, the MG and $\boldsymbol{b}$ are then the same as if $t \leq \min(r_1, r_2)$ and are given by Theorem 3.

Proof: A detailed proof is given in [13] but the idea behind the proof is very simple. It comes simply from observing which of the three constraints $b_1 \leq r_1$, $b_2 \leq r_2$, and $b_1 + b_2 \leq t$ is active. This will be clear after the following discussion. ■

Geometrical insight: In Fig. 1, the MG region is plotted for all the stream allocations in the system setting $\boldsymbol{r} = (6, 6)$ and $t = 6$. Since $t \leq \min(r_1, r_2)$, only the constraint $b_1 + b_2 = t$ is active and Theorem 3 can be applied. We can also observe the dashed lines which represent the MG with a real number of transmitted streams. The common points between this line and the rectangles correspond to the rate coefficients $\gamma_{1,\max}(\cdot)$. Each rectangle is associated with a stream allocation and we only need to find the intersection points to obtain the rate region and all the results of Theorem 3.

We have marked the MG region for the antenna configuration $\boldsymbol{r} = (2, 4)$, in which case $t \geq r_1 + r_2$ and the constraint $b_1 + b_2 \leq t$ is not active, such that Theorem 2 holds.

Finally, if $\boldsymbol{r} = (4, 4)$, it can be observed in Fig. 1 that each constraint is active on some part of the rate region. We have indicated the rate coefficients ϕ_1 and ϕ_2 from Theorem 4 and emphasized the boundary of the MG region. It becomes then clear how Theorem 3 can be applied between ϕ_1 and ϕ_2, while it is otherwise Theorem 2 which holds.

IV. Rate Offset Optimization

The case where at least one user applies FM is studied in [10], such that we now focus on the case when no user applies FM and the constraint $b_1 + b_2 = t$ is active. The results are then

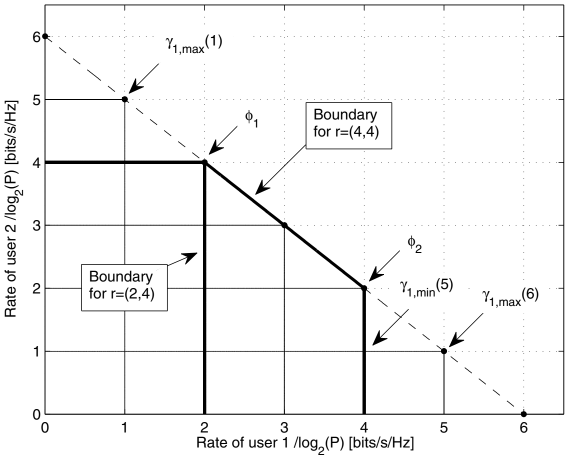

Fig. 1. Rate region divided by $\log(P)$ when P tends to infinity with $t = 6$ for the different number of streams transmitted.

dependent on whether $b_1+b_2 \leq r_i$ or not for each user. We will consider for simplicity the two cases $b_1+b_2 \leq \min(r_1, r_2)$ and $b_1 + b_2 \geq \max(r_1, r_2)$, from which the general configuration follows trivially. Since we consider the domain of the rate region where $b_1 + b_2 = t$, the differentiation of the two cases depends only on the antenna configuration.

A. When $b_1 + b_2 \leq \min(r_1, r_2)$

Theorem 5. *Let $b_1 + b_2 \leq \min(r_1, r_2)$.*

If $\gamma_1 \in]\gamma_{1,\min}(i-1), \gamma_{1,\max}(i)[, i \in \{1, \ldots, t-1\}$, the asymptotically optimal NP matrices are the ones maximizing the rate of user 2, denoted as $(\bar{\mathbf{T}}_{1,\emptyset_2}, \bar{\mathbf{T}}_{2,\emptyset_2})$, and given as the b_2 eigenvectors of the matrix $\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2$, and the b_1 eigenvectors of the matrix $\mathbf{P}_{\perp,1}\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1\mathbf{P}_{\perp,1}$, respectively, where

$$\mathbf{P}_{\perp,1} := \mathbf{I} - \mathbf{H}_1^{\mathrm{H}}\mathbf{H}_2\bar{\mathbf{T}}_{2,\emptyset_2}(\bar{\mathbf{T}}_{2,\emptyset_2}^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_1\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_2\bar{\mathbf{T}}_{2,\emptyset_2})^{-1}\bar{\mathbf{T}}_{2,\emptyset_2}^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_1,$$

The asymptotic sum rate then reads as

$$R_\infty(\boldsymbol{\gamma}) = \frac{b_2}{\gamma_2}\left(\log\left(\frac{P}{b_2}\right) + \frac{1}{b_2}\sum_{\ell=1}^{b_2}\log(\lambda_\ell(\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2))\right),$$

with $\lambda_\ell(\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2)$ the ℓ-th largest eigenvalue of the Gramian $\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2$ of the channel of user 2.

If $\gamma_1 \in]\gamma_{1,\max}(i), \gamma_{1,\min}(i+1)[, i \in \{1, \ldots, t-1\}$, the asymptotically optimal NP matrices are the ones maximizing the rate of user 1, denoted as $(\bar{\mathbf{T}}_{1,\emptyset_1}, \bar{\mathbf{T}}_{2,\emptyset_1})$, and defined as the b_1 eigenvectors of the matrix $\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1$, and the b_2 eigenvectors of the matrix $\mathbf{P}_{\perp,2}\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2\mathbf{P}_{\perp,2}$, respectively, where

$$\mathbf{P}_{\perp,2} := \mathbf{I} - \mathbf{H}_2^{\mathrm{H}}\mathbf{H}_1\bar{\mathbf{T}}_{1,\emptyset_1}(\bar{\mathbf{T}}_{1,\emptyset_1}^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_2\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_1\bar{\mathbf{T}}_{1,\emptyset_1})^{-1}\bar{\mathbf{T}}_{1,\emptyset_1}^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_2.$$

The asymptotic sum rate then reads as

$$R_\infty(\boldsymbol{\gamma}) = \frac{b_1}{\gamma_1}\left(\log\left(\frac{P}{b_1}\right) + \frac{1}{b_1}\sum_{\ell=1}^{b_1}\log(\lambda_\ell(\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1))\right),$$

with $\lambda_\ell(\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1)$ the ℓ-th largest eigenvalue of the Gramian $\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1$ of the channel of user 1.

Proof: We denote the limiting user as user ℓ and the other user as user $n\ell$. The approximated rate of user ℓ then reads as

$$R_\ell'' = b_\ell\left(\log(P - P_{n\ell}) - \log(b_\ell) - \log(c_\ell(\bar{\mathbf{T}}_1, \bar{\mathbf{T}}_2))\right).$$

From the scaling in P of the power allocation of the two users in Theorem 1, we can deduce that the non-limiting user is allocated with a negligible fraction of P when P tends to infinity. Thus, the term $1 - P_{n\ell}/P$ tends to 1, and since the rate ratio constraint is already fulfilled by the power allocation, it is only necessary to optimize $c_\ell(\bar{\mathbf{T}}_1, \bar{\mathbf{T}}_2)$. This means that the asymptotically optimal NP matrices are the ones maximizing the rate of the limiting user.

The NP matrix $\bar{\mathbf{T}}_{n\ell}$ has to be chosen in order to maximize the quotient of determinants in the expression of c_ℓ in (4). Decomposing the Gramian of the projected channel shows that the maximal value of the ratio of determinants in (4) is 1 and that this value is achieved if and only if the two users emit orthogonally to each other, i.e., $\bar{\mathbf{T}}_1^{\mathrm{H}}\bar{\mathbf{H}}_1^{\mathrm{H}}\bar{\mathbf{H}}_2\bar{\mathbf{T}}_2 = \mathbf{0}_{b_1 \times b_2}$. $\bar{\mathbf{T}}_{n\ell}$ has to be in the orthogonal complement of $\bar{\mathbf{T}}_\ell^{\mathrm{H}}\bar{\mathbf{H}}_\ell^{\mathrm{H}}\bar{\mathbf{H}}_{n\ell}$ of size $b_\ell \times r_{n\ell}$, thus of dimension $r_{n\ell} - b_\ell$. Since $b_1 + b_2 \leq r_{n\ell}$, this condition can be fullfilled. Thus, $\bar{\mathbf{T}}_{n\ell}$ is set as in the theorem and the expressions for $\bar{\mathbf{T}}_\ell$ follows trivially. ■

B. When $b_1 + b_2 > \max(r_1, r_2)$

We can observe from the proof of Theorem 5 that the condition to be able to apply the theorem over the domain where user 1 is the limiting user is $b_1 + b_2 \leq r_2$, and $b_1 + b_2 \leq r_1$ when user 2 is the limiting user. We now study for clarity the case where $b_1 + b_2 > \max(r_1, r_2)$, but it is straightforward that if we have $r_1 < t < r_2$, for example, then Theorem 5 will apply when user 1 is the limiting user, and otherwise Theorem 6, proven in the following.

The intuitive explanation for the differences between the two cases is that when $b_1 + b_2 > r_{n\ell}$, it is in general not possible to let the non-limiting user transmit without creating any interference to the limiting user such that the NP matrices depend on each other and an iterative algorithm is needed.

Lemma 1. *Let $\left(\bar{\mathbf{T}}_1^{(n)}, \bar{\mathbf{T}}_2^{(n)}\right)$ be given at step n. Choosing $\bar{\mathbf{T}}_2^{(n+1)}$ as the b_2 principal eigenvectors of the generalized eigenvalue problem $\left(\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2, \bar{\mathbf{H}}_2^{\mathrm{H}}(\bar{\mathbf{T}}_1^{(n)})\bar{\mathbf{H}}_2(\bar{\mathbf{T}}_1^{(n)})\right)$, and then $\bar{\mathbf{T}}_1^{(n+1)}$ as the b_1 principal eigenvectors of $\bar{\mathbf{H}}_1^{\mathrm{H}}(\bar{\mathbf{T}}_2^{(n+1)})\bar{\mathbf{H}}_1(\bar{\mathbf{T}}_2^{(n+1)})$, yields a sequence of matrices converging almost surely to local maximizers of the rate shift of user 1, denoted as $(\bar{\mathbf{T}}_{1,\mathrm{alg}_1}, \bar{\mathbf{T}}_{2,\mathrm{alg}_1})$.*

Similarly, choosing $\bar{\mathbf{T}}_1^{(n+1)}$ as the b_1 principal eigenvectors of the generalized eigenvalue problem $\left(\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1, \bar{\mathbf{H}}_1^{\mathrm{H}}(\bar{\mathbf{T}}_2^{(n)})\bar{\mathbf{H}}_1(\bar{\mathbf{T}}_2^{(n)})\right)$, and then $\bar{\mathbf{T}}_2^{(n+1)}$ as the b_2 principal eigenvectors of $\bar{\mathbf{H}}_2^{\mathrm{H}}(\bar{\mathbf{T}}_1^{(n+1)})\bar{\mathbf{H}}_2(\bar{\mathbf{T}}_1^{(n+1)})$, yields a sequence of matrices converging almost surely to local maximizers of the rate shift of user 2, denoted as $(\bar{\mathbf{T}}_{1,\mathrm{alg}_2}, \bar{\mathbf{T}}_{2,\mathrm{alg}_2})$.

Proof: We now consider w.l.o.g. the first part of the lemma corresponding to the maximization of the rate shift of user 1. Two formulations for the rate shift of user 1 are:

$$c_1(\bar{\mathbf{T}}_1,\bar{\mathbf{T}}_2)=\left|\bar{\mathbf{T}}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1\bar{\mathbf{T}}_1\right|^{-\frac{1}{b_1}}\frac{\left|\bar{\mathbf{T}}_2^{\mathrm{H}}\bar{\mathbf{H}}_2^{\mathrm{H}}(\bar{\mathbf{T}}_1)\bar{\mathbf{H}}_2(\bar{\mathbf{T}}_1)\bar{\mathbf{T}}_2\right|^{-\frac{1}{b_1}}}{\left|\bar{\mathbf{T}}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2\bar{\mathbf{T}}_2\right|^{-\frac{1}{b_1}}}, \tag{12}$$

$$c_1(\bar{\mathbf{T}}_1,\bar{\mathbf{T}}_2)=\left|\bar{\mathbf{T}}_1^{\mathrm{H}}\bar{\mathbf{H}}_1^{\mathrm{H}}(\bar{\mathbf{T}}_2)\bar{\mathbf{H}}_1(\bar{\mathbf{T}}_2)\bar{\mathbf{T}}_1\right|^{-\frac{1}{b_1}}. \tag{13}$$

From (12), we observe that the update of $\bar{\mathbf{T}}_2$ leads to the global maximum for given $\bar{\mathbf{T}}_1^{(n)}$, and from (13), the same holds when updating $\bar{\mathbf{T}}_1$ for given $\bar{\mathbf{T}}_2^{(n+1)}$. Since we obtain at every step the global maximum, the objective increases monotonically. It is clearly upper bounded, and hence converges monotonically to an optimum, which is almost surely a local maximum. ∎

Theorem 6. *Let $b_1+b_2>\max(r_1,r_2)$, and denote as $(\bar{\mathbf{T}}_{1,\varnothing_1},\bar{\mathbf{T}}_{2,\varnothing_1})$ and $(\bar{\mathbf{T}}_{1,\varnothing_2},\bar{\mathbf{T}}_{2,\varnothing_2})$, the NP matrices, a priori unknown, maximizing the rate shift of user 1 and 2, respectively. For $i\in\{1,\dots,t-1\}$, it holds:*

If $\gamma_1\in]\gamma_{1,\max}(i),\gamma_{1,\min}(i+1)[$, $R_\infty(\boldsymbol{\gamma})$ then reads as

$$R_\infty(\boldsymbol{\gamma})=\frac{b_1}{\gamma_1}\left(\log\left(\frac{P}{b_1}\right)-\log(c_1(\bar{\mathbf{T}}_{1,\varnothing_1},\bar{\mathbf{T}}_{2,\varnothing_1}))\right),$$

If $\gamma_1\in]\gamma_{1,\min}(i),\gamma_{1,\max}(i)[$, $R_\infty(\boldsymbol{\gamma})$ then reads as

$$R_\infty(\boldsymbol{\gamma})=\frac{b_2}{\gamma_2}\left(\log\left(\frac{P}{b_2}\right)-\log(c_2(\bar{\mathbf{T}}_{1,\varnothing_2},\bar{\mathbf{T}}_{2,\varnothing_2}))\right),$$

Proof: The proof follows exactly the same steps as the proof of Theorem 5. ∎

The optimal NP matrices used in Theorem 6 are not known, but Lemma 1 gives NP matrices maximizing locally $R_\infty(\boldsymbol{\gamma})$. If these NP matrices are also global maximizer, they can be used directly in Theorem 6. Otherwise, only a lower bound for the asymptotic sum rate is obtained.

V. Rate Region at Finite SNR

In section IV, we have derived the asymptotic sum rate and we now want to obtain results valid at finite SNR. The optimal NP matrices are much more difficult to derive at finite SNR due to the dependency between the power allocation and the NP matrices. However, once the NP matrices are fixed, finding the power allocation fulfilling the rate ratio constraint (9) can be done very easily via convex algorithms or by solving only approximately the rate ratio constraint [12], [13].

It is then possible by fixing the rate shifts first and then calculating the power allocation to derive lower and upper bounds for the sum rate, or equivalently inner and outer bounds for the rate region boundary, at high but finite SNR. Indeed, in Section IV, we have derived NP matrices optimizing the rate of only one of the users which will clearly lead to lower bounds. Furthermore, when $b_1+b_2\le\min(r_1,r_2)$, the optimal NP matrices for one user are reached. It means that by using the rate shifts $(c_{1,\mathrm{v}},c_{2,\mathrm{v}}):=(c_1(\bar{\mathbf{T}}_{1,\varnothing 2},\bar{\mathbf{T}}_{2,\varnothing 2}),c_2(\bar{\mathbf{T}}_{1,\varnothing 2},\bar{\mathbf{T}}_{2,\varnothing 2}))$ we obtain an outer bound for the rate region boundary when $b_1+b_2\le t$, since both users use their optimal rate shifts.

When $b_1+b_2>t$, two inner bounds can also be obtained by using $(\bar{\mathbf{T}}_{1,\mathrm{alg}_1},\bar{\mathbf{T}}_{2,\mathrm{alg}_1})$ and $(\bar{\mathbf{T}}_{1,\mathrm{alg}_2},\bar{\mathbf{T}}_{2,\mathrm{alg}_2})$, respectively. However, the lower bounds are not in closed form but obtained with an algorithm and the outer bound cannot be derived following the same method, since these NP matrices are only local maximizer and not global maximizer. Nevertheless, it is easy to derive some other suboptimal NP matrices in closed form (by choosing the NP matrix of one user arbitrarily, for example) and a loser closed form outer bound can also be obtained by assuming that the two users do not interfere with each other. These bounds are not further presented here due to space constraint, but the derivations are straightforward.

Finally, when both users transmit only one stream, which is called *beamforming* (BF), it is possible to derive an algorithm converging this time not to a bound but to a local maximum of the sum rate. The method used is the fixed coordinate (FC) approach from [9], which consists in fixing the rate of user 1 to the constant $\log(c)$, and then maximizing the rate of user 2. The constraint $R_1''=\log(c)$ clearly leads to

$$P_1=\frac{c}{\bar{\boldsymbol{t}}_1^{\mathrm{H}}\bar{\mathbf{H}}_1^{\mathrm{H}}(\bar{\boldsymbol{t}}_2)\bar{\mathbf{H}}_1(\bar{\boldsymbol{t}}_2)\bar{\boldsymbol{t}}_1}, \tag{14}$$

where $\boldsymbol{t}_1$ and $\boldsymbol{t}_2$ are the NP matrices of user 1 and 2, respectively, when they both apply BF. The power allocation (14) can then be inserted in the rate of user 2 to yield

$$R_2''=\log\left(\left(\frac{\bar{\boldsymbol{t}}_1^{\mathrm{H}}\left(P\bar{\mathbf{H}}_1^{\mathrm{H}}(\bar{\boldsymbol{t}}_2)\bar{\mathbf{H}}_1(\bar{\boldsymbol{t}}_2)-c\mathbf{I}_{r_1}\right)\bar{\boldsymbol{t}}_1}{\bar{\boldsymbol{t}}_1^{\mathrm{H}}\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1\bar{\boldsymbol{t}}_1}\right)\bar{\boldsymbol{t}}_2^{\mathrm{H}}\mathbf{H}_2^{\mathrm{H}}\mathbf{H}_2\bar{\boldsymbol{t}}_2\right), \tag{15}$$

$$R_2''=\log\left(\bar{\boldsymbol{t}}_2^{\mathrm{H}}\left(P\bar{\mathbf{H}}_2^{\mathrm{H}}(\bar{\boldsymbol{t}}_1)\bar{\mathbf{H}}_2(\bar{\boldsymbol{t}}_1)-\frac{c}{\bar{\boldsymbol{t}}_1^{\mathrm{H}}\bar{\mathbf{H}}_1^{\mathrm{H}}\bar{\mathbf{H}}_1\bar{\boldsymbol{t}}_1}\mathbf{I}_{r_2}\right)\bar{\boldsymbol{t}}_2\right). \tag{16}$$

Theorem 7. *Let $(\bar{\boldsymbol{t}}_1^{(n)},\bar{\boldsymbol{t}}_2^{(n)})$ be given at step n, setting $\bar{\boldsymbol{t}}_1^{(n+1)}$ as the principal eigenvector of the generalized eigenvalue problem $(\mathbf{H}_1^{\mathrm{H}}\mathbf{H}_1,P\bar{\mathbf{H}}_1^{\mathrm{H}}(\bar{\boldsymbol{t}}_2)\bar{\mathbf{H}}_1(\bar{\boldsymbol{t}}_2)-c\mathbf{I}_{r_1})$, and then $\bar{\boldsymbol{t}}_2^{(n+1)}$ as the principal eigenvector of $P\bar{\mathbf{H}}_2^{\mathrm{H}}(\bar{\boldsymbol{t}}_1^{(n+1)})\bar{\mathbf{H}}_2(\bar{\boldsymbol{t}}_1^{(n+1)})-c/(\bar{\boldsymbol{t}}_1^{(n+1)\mathrm{H}}\bar{\mathbf{H}}_1^{\mathrm{H}}\bar{\mathbf{H}}_1\bar{\boldsymbol{t}}_1^{(n+1)})\mathbf{I}_{r_2}$ yields a sequence of BF vectors converging almost surely to local maximizers of the sum rate, denoted as $(\bar{\boldsymbol{t}}_{1,\mathrm{alg}},\bar{\boldsymbol{t}}_{2,\mathrm{alg}})$.*

Proof: The proof follows the same method as the proof of Theorem 5 using the two expressions (15) and (16). ∎

VI. Simulations

In Fig. 2, we consider the antenna configuration $\boldsymbol{r}=(4,4)$, with $t=4$ and $P=40$ dB, when both users apply BF. We consider this stream allocation because it behaves like any other stream allocations fulfilling $b_1+b_2\le\min(r_1,r_2)$. The only difference is that the algorithmic solution from Theorem 7 can be used and gives a verification of the accuracy of the bounds. We have also plotted with the dashed line the asymptotic rate region derived in Section IV.

First, the high-SNR approximated rate region at finite SNR is similar to the asymptotic rate region, only with the corner "rounded-down". Second, the lower bound, in which the rate shift of the limiting user is optimized, is very close to the

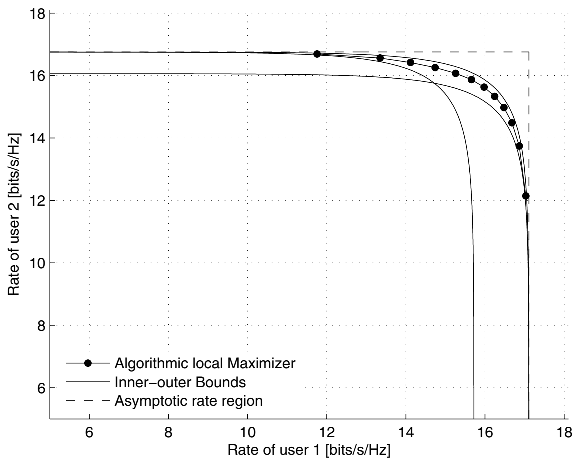

Fig. 2. Approximate rate region with $\boldsymbol{r} = (4, 4)$, $t = 4$, and $P = 40$ dB when both users apply BF.

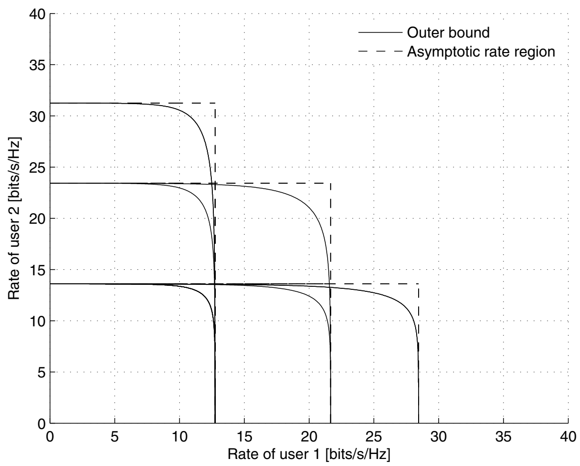

Fig. 3. Outer bounds of the approximate rate region with $\boldsymbol{r} = (4, 4)$, $t = 4$, and $P = 40$ dB for different number of transmitted streams.

outer bound, and both are tight over large domains of the rate region. Finally, the algorithmic solution is always between the two bounds and hence converges to the global optimum at least in the region where the bounds are tight.

In Fig. 3, the outer bound is plotted for all stream allocations when $\boldsymbol{r} = (4, 4)$, $t = 4$ and $P = 40$ dB. For clarity, we have kept only the outer bound because it is tight over a large part of the rate region and the behavior of the inner bounds is clear from Fig. 2. We have also plotted the asymptotic rate region derived in Section IV. We can observe the good fit between the asymptotic rate region and the outer bound at finite SNR. Moreover, the intersections between the different stream allocations occur at practically the same points for the outer bound and the asymptotic rate region. Since the asymptotic rate region is known in closed form, we can obtain closed form expressions approximating accurately the intersection points between the different stream allocations.

We have plotted in Fig. 2 and Fig. 3 the high-SNR approximated rate region obtained using the rate expressions (2). Since the approximation from (1) to (2) consists in neglecting the identities, the inner bound is also valid for the exact rate region boundary. However, the outer bound is valid only when the high-SNR approximation error is negligible. The error has been shown in [13] to be significant only close to the axes, where it is optimal to let one user apply BF. An improvement of the high-SNR approximation when one user applies BF is given in [12], [13] and leads to a very accurate approximation of the integrality of the rate region, for P as low as 30 dB.

VII. Conclusion

We have studied the rate region at high-SNR with linear precoding when the transmitter has fewer antennas than the sum of the antennas at the receivers. The asymptotic rate region has been derived in closed form and we have shown that the precoding matrices derived can be used to obtain accurate inner and outer bounds for the rate region at finite SNR.

Our asymptotic results are given in closed forms as functions of the eigenvalues of the Gramian of the channels, and can thus be easily evaluated in common fading scenarios (uncorrelated Rician, correlated Rayleigh). The approach is geometric and intuitive such that it gives a good insight into the rate region at finite SNR. Finally, the approach presented has a very good potential to be extended to the K user case.

References

[1] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of Gaussian MIMO broadcast channels," *IEEE Transactions on Information Theory*, vol. 49, no. 10, october 2003.

[2] H. Weingarten, S. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Transactions on Information Theory*, vol. 52, no. 9, september 2006.

[3] R. Böhnke and K.-D. Kammeyer, "Weighted sum rate maximization for the MIMO-downlink using a projected conjugate gradient algorithm," in *Proc. IWCLD*, 2007.

[4] U. Erez and S. ten Brink, "A close-to-capacity dirty paper coding scheme," *IEEE Transactions on Information Theory*, vol. 51, no. 10, october 2005.

[5] R. Hunger, D. Schmidt, and M. Joham, "A combinatorial approach to maximizing the sum rate in the MIMO BC with linear precoding," in *Proc. Asilomar*, 2008.

[6] S. Shi, M. Schubert, and H. Boche, "Rate optimization for multiuser MIMO systems with linear processing," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, Aug. 2008.

[7] J. Lee and N. Jindal, "High SNR analysis for MIMO broadcast channels: dirty paper coding versus linear precoding," *IEEE Transactions on Information Theory*, vol. 53, no. 12, december 2007.

[8] R. Hunger and M. Joham, "An asymptotic analysis of the MIMO broadcast channel under linear filtering," in *Proc. CISS*, 2009.

[9] R. Hunger, P. de Kerret, and M. Joham, "The geometry of the MIMO broadcast channel rate region under linear filtering at high SNR," in *Proc. Asilomar*, 2009.

[10] P. de Kerret, R. Hunger, M. Joham, W. Utschick, and R. Mathar, "MIMO Broadcast Channel Rate Region with Linear Precoding at High SNR with Full Multiplexing," submitted to Allerton conference.

[11] R. Hunger and M. Joham, "A general rate duality of the MIMO multiple access channel and the MIMO broadcast channel," in *Proc. GLOBECOM*, 2008.

[12] P. de Kerret, "The geometry of the high-power MIMO BC rate region under linear filtering," Master's thesis, Technische Universität München (TUM), 2009.

[13] P. de Kerret, R. Hunger, M. Joham, W. Utschick, and R. Mathar, "Analysis of the two user MIMO-BC rate region with linear precoding at high SNR," In preparation.