Signaling over MIMO Multi-Base Systems: Combination of Multi-Access and Broadcast Schemes

Mohammad Ali Maddah-Ali, Abolfazl S. Motahari, and Amir K. Khandani
Coding & Signal Transmission Laboratory (www.cst.uwaterloo.ca)
Dept. of Elec. and Comp. Eng., University of Waterloo
Waterloo, ON, Canada, N2L 3G1
Tel: 519-725-7338, Fax: 519-888-4338
e-mail: {mohammad, abolfazl khandani}@cst.uwaterloo.ca

***Abstract*— A new structure for multi-base systems is studied in which each user receives data from two nearby base stations, rather than only from the strongest one. This system can be considered as a combination of broadcast and multi-access channels. By taking advantages of both perspectives, an achievable rate region for a discrete memoryless channel modeled by $\Pr(y_1, y_2|x_1, x_2)$ is derived. In this model, x_1 and x_2 represent the transmitted signals by the transmitter one and two, respectively, and y_1 and y_2 denote the received signals by the receiver one and two, respectively. In this derivation, it is assumed that each transmitter is unaware of the data of the other transmitter, and therefore x_1 and x_2 are independent. To investigate the advantage of this scheme, an efficient signaling method which works at a corner point of the achievable region for multiple-antenna scenarios is developed. In the proposed scheme, each base station only requires the state information of the channels between the other base station and each user. In this paper, the signaling scheme is elaborated for the case that each transmitter/reciever is equipped with three antennas. It is proven that in such a scenario, the multiplexing gain of four is achievable, which outperforms any other conventional schemes.**

I. Introduction

In conventional wireless systems, each user receives information from one base station, which is generally the strongest one. In this case, the performance of the system can be dramatically deteriorated by the interference from the other pairs of transmitters/recivers, known as *co-channel interference*.

A number of research works have investigated the effect of co-channel interference in multi-input multi-output (MIMO) multi-user systems. In [1], the capacity of system for a group of interfering users employing single-user detection is studied. In [2], [3], multi-user detection and turbo decoding are exploited to improve the performance of the system, where it is assumed that all users have complete information of their channels with the base stations. A well-known metric to evaluate the performance of the signaling schemes is the so-called *multiplexing gain* which is defined as the ratio of the sum-rate over log(SNR) for large signal-to-noise-ratio (SNR) values. Simulation results indicate that the multiplexing gain of the signaling scheme proposed in [2] is zero. In other words, the sum-rate converges to a fixed value as SNR increases. In fact, the un-canceled terms of co-channel interference in the denominator of the signal-to-interference-plus-noise-ratio (SINR) do not allow further increase in the rate. In [4], the multiplexing gains of MIMO multiuser schemes are investigated. Specially, it is proven that in an interference channel with two transmitters, and two receivers, where each of them is equipped with η antennas, the conventional signaling schemes can achieve the maximum multiplexing gain is η.

To mitigate the co-channel interference, the cooperation among base stations is proposed [5]. In [5], the infinite-capacity link among base stations is assumed which reduces the system to a single broadcast channel. The QR decomposition scheme as a signaling method over MIMO broadcast channel is applied which completely cancels the interference. The interference cancelation is based on a result known as dirty paper coding (DPC) due to Costa, 1983 [6]. In [7], the performance of the method proposed in [5] is evaluated for a more practical channel model. In [8], the idea of [5] is explored taking into account the individual power constraints per base station and making use of uplink–downlink duality. Cooperative base stations in uplink is considered in [9]. By assuming a full cooperation among base stations, the system is simplified to a single MIMO multi-access system.

In this paper, a new signaling scheme for multi-base systems in downlink is proposed. In this scheme, each user receives data from two nearby base stations, rather than only from the closest one. In this case, we can consider the system as a set of broadcast channels (from base stations' point of view) or a set of multi-access channels (from users' point of view). We benefit from both perspectives to derive an achievable rate region for a discrete memoryless channel modeled by $\Pr(y_1, y_2|x_1, x_2)$. In this derivation, it is assumed that each transmitter is unaware of the data of the other transmitter, and therefore x_1 and x_2 are independent. This derivation is based on a combination of two achievable regions: (i) Marton rate region for the memoryless broadcast channels [10], [11], and (ii) rate region for the memoryless multi-access channels [12]. By focusing on a corner point of the derived achievable region, an efficient signalling scheme for such systems in proposed. In the proposed scheme, each base station only requires the state information of the channels between the other base station and each user. In this paper, the proposed signalling scheme is elaborated for the case that each of the transmitters and receivers is equipped with three antennas. It is proven that in such a scenario, the multiplexing gain of four is achievable, which outperforms any other conventional schemes.

II. Achievable Region

In the following, an achievable rate region for a general discrete memoryless channel, modeled by $\Pr(y_1, y_2|x_1, x_2)$, is derived. In the suggested rate region, the auxiliary random variables W_1 and Z_1 contain information from the transmitter one to the receivers one and two, respectively. Similarly, the auxiliary random variables W_2 and Z_2 contain information from the transmitter two to the receivers two and one, respectively (see Fig. 1).

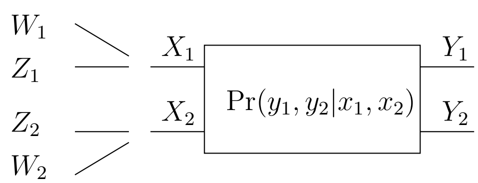

Fig. 1. A General Discrete Memoryless Channel

Theorem 1 *Consider a discrete memoryless channel modelled by* $\Pr(y_1, y_2|x_1, x_2)$. *Assume that the transmitter* t, $t = 1, 2$, *transmits data to the receiver* r, $r = 1, 2$, *with the rate* R_{rt}. *Then, an achievable region is given by the set of all rates in the convex closure of the quadruples* $(R_{11}, R_{21}, R_{12}, R_{22})$ *satisfying,*

$$R_{11} \le q_{11} \le I(Y_1; W_1|Z_2), \tag{1}$$
$$R_{21} \le q_{21} \le I(Y_2; Z_1|W_2), \tag{2}$$
$$R_{22} \le q_{22} \le I(Y_2; W_2|Z_1), \tag{3}$$
$$R_{12} \le q_{12} \le I(Y_1; Z_2|W_1), \tag{4}$$
$$R_{11} + R_{21} \le q_{11} + q_{21} - I(Z_1; W_1), \tag{5}$$
$$R_{22} + R_{12} \le q_{22} + q_{12} - I(Z_2; W_2), \tag{6}$$
$$q_{11} + q_{12} \le I(Y_1; W_1, Z_2), \tag{7}$$
$$q_{22} + q_{21} \le I(Y_2; W_2, Z_1), \tag{8}$$

for some joint distribution of $\Pr(w_1, z_1, x_1, w_2, z_2, x_2,) = \Pr(w_1, z_1, x_1) \Pr(w_2, z_2, x_2)$. *In addition, by defining* $R_r = R_{r1} + R_{r2}$, $r = 1, 2$, *as the total rate of the receiver* r, *an achievable rate region for* R_1 *and* R_2 *is given by convex closure of all rates* (R_1, R_2), *satisfying*

$$\begin{aligned} &R_1 \le I(Y_1; Z_2, W_1), \;\; R_2 \le I(Y_2; Z_1, W_2) \\ &R_1 + R_2 \le I(Y_1; Z_2, W_1) + I(Y_2; Z_1, W_2) \\ &-I(Z_1; W_1) - I(Z_2; W_2) \end{aligned}$$

for some joint distribution as in the first part.

Proof: Fix $\Pr(w_1, z_1, x_1, w_2, z_2, x_2,) = \Pr(w_1, z_1, x_1) \Pr(w_2, z_2, x_2)$.

Random Encoding for the Transmitter One: Generate $2^{nq_{11}}$ i.i.d. sequences $w_1^n \in A_\epsilon^n$ according to the uniform distribution over $A_\epsilon^n(W_1)$, where $A_\epsilon^n(W_1)$ denotes the typical set for the random variable Z_1^n. Label the selected i.i.d. sequences as $w_1^n(k_1)$, $k_1 = 1, \dots, 2^{nq_{11}}$. Similarly, generate $2^{nq_{21}}$ i.i.d. sequences $z_1^n \in A_\epsilon^n$ according to the uniform distribution over $A_\epsilon^n(Z_1)$, where $A_\epsilon^n(Z_1)$ denotes the typical set for the random variable Z_1^n. Label the selected i.i.d. sequences as $z_1^n(l_1)$, $l_1 = 1, \dots, 2^{nq_{21}}$. For $i_1 \in [1, 2^{nR_{11}}]$ and $j_1 \in [1, 2^{nR_{21}}]$, define the cells

$$\begin{aligned} &B_{i_1}^{(1)} = \Big\{ w_1^n(k_1) : \\ &k_1 \in [(i_1 - 1)2^{n(q_{11} - R_{11})} + 1, i_1 2^{n(q_{11} - R_{11}) + 1}] \Big\}, \\ &C_{j_1}^{(1)} = \Big\{ z_1^n(l_1) : \\ &l_1 \in [(j_1 - 1)2^{n(q_{21} - R_{21})} + 1, j_1 2^{n(q_{21} - R_{21}) + 1}] \Big\}, \\ &D_{i_1 j_1}^{(1)} = \Big\{ (w_1^n(k_1), z_1^n(l_1)) : \\ &w_1^n(k_1) \in B_{i_1}^1, z_1^n(l_1) \in C_{j_1}^1, (w_1^n(k_1), z_1^n(l_1)) \in A_\epsilon^n \Big\}. \end{aligned}$$

To send a message pair (i_1, j_1), choose one pair $(w_1^n(k_1), z_1^n(l_1))$ from $D_{i_1 j_1}^{(1)}$, and find an $x_1^n(i_1, j_1)$ that is jointly typical with that pair.

Random Encoding for the Transmitter Two: Similarly, for the user two, generate $w_2(k_2)$ for $1 \le k_2 \le 2^{nq_{22}}$, $z_2(l_2)$ for $1 \le l_2 \le 2^{nq_{12}}$, and cells $B_{i_2}^{(2)}$, $C_{j_2}^{(2)}$, $D_{i_2 j_2}^{(2)}$, for $1 \le i_2 \le 2^{nR_{22}}$, and $1 \le j_2 \le 2^{nR_{12}}$. For message pair (i_2, j_2), choose one pair $(w_2^n(k_2), z_2^n(l_2))$ from $D_{i_2 j_2}^{(2)}$, and find an $x_2^n(i_2, j_2)$ that is jointly typical with that pair.

Decoding: Receiver one finds the unique indices pair (k_1, l_2) such that $(w_1^n(k_1), z_2^n(l_2), y_1^n) \in A_\epsilon^n$. Similarly, receiver two finds the unique indices pair (k_2, l_1) such that $(w_2^n(k_2), z_1^n(l_1), y_2^n) \in A_\epsilon^n$.

Using the above random coding scheme, we can prove that the average probability of error converges to zero as $n \longrightarrow \infty$, if inequalities (1) to (8) are satisfied. The second part of the theorem is derived directly from the first part. ■

A. A Corner Point

To show the advantages of this scheme, we focus on one of the corner points of the achievable region for the rate vector $(R_{11}, R_{12}, R_{21}, R_{22})$, for a fixed joint probability on inputs and auxiliary random variables. To this end, we choose for R_{11} and R_{22} the maximum possible values, i.e.

$$R_{11} = I(Y_1; W_1|Z_2), \tag{9}$$
$$R_{22} = I(Y_2; W_2|Z_1). \tag{10}$$

With these choices of R_{11} and R_{22}, we can show that the maximum possible values for R_{12} and R_{21} are equal to,

$$R_{12} = I(Y_1; Z_2) - I(Z_2; W_2), \tag{11}$$
$$R_{21} = I(Y_2; Z_1) - I(Z_1; W_1). \tag{12}$$

Here, we investigate the rates of the data received by the user 1, i.e. R_{11} and R_{12}. Equation (9) implies that to achieve the highest rate for W_1, receiver one first decodes Z_2, and then W_1. The formula obtained for R_{12} is basically the same as that of the rate of the channel with non-causally known i.i.d. state at the transmitter, derived by Gelfand-Pinsker [13]. In fact, if the transmitter two first chooses a codeword for W_2, its interference over Z_2 at the receiver one terminal is non-causally known by the transmitter, and therefore rate of (11) is achievable. For the special case of additive white Gaussian noise, with Gaussian distribution for auxiliary random variables W_2 and Z_1, equation (11) implies that the interference of W_2 over Z_1 at the receiver one terminal can be effectively canceled out [6]. This result is known as the *dirty paper coding*, due to Costa [6].

The above observation leads us to a signaling scheme which is elaborated for MIMO scenarios in the next section.

III. Signaling Method for MIMO Systems

Consider a MIMO multi-base system with the base stations t, $t=1,2$, as the transmitters and the users r, $r=1,2$ as the receivers. As an example, here we focus on a case where each base station t is equipped with $M_t=3$ antennas, and similarly each user r, $r=1,2$, is equipped with $N_r=3$ antennas. However, the proposed scenario can be generalized to the case of different number of antennas. Assuming flat fading environment, the channel between the base station t and the user r is represented by the channel matrix $\mathbf{H}_{rt}$, where $\mathbf{H}_{rt} \in \mathcal{C}^{3\times 3}$. The received vector $\mathbf{y}_r \in \mathcal{C}^{3\times 1}$ by user r, $r=1,2$, is given by,

$$\mathbf{y}_1 = \mathbf{H}_{11}\mathbf{x}_1 + \mathbf{H}_{12}\mathbf{x}_2 + \mathbf{n}_1, \tag{13}$$

$$\mathbf{y}_2 = \mathbf{H}_{21}\mathbf{x}_1 + \mathbf{H}_{22}\mathbf{x}_2 + \mathbf{n}_2, \tag{14}$$

where $\mathbf{x}_t \in \mathcal{C}^{3\times 1}$ represents the transmitted vector by the base station t. The vector $\mathbf{n}_r \in \mathcal{C}^{3\times 1}$ is a white Gaussian noise with zero mean and identity covariance matrix. It is assumed that $E(\mathbf{x}_t\mathbf{x}_t^{\dagger}) \le P$, for $t=1,2$.

In the proposed scenario, each base station transmits two data streams. The base station t sends the data stream $d_{1,t}$ to the user 1 and the data stream $d_{2,t}$ to the user 2. The transmitted vectors are equal to the linear superposition of the modulation vectors with d_{rt}, $t,r=1,2$, as the coefficients, i.e.

$$\mathbf{x}_1 = d_{11}\mathbf{v}_{11} + d_{21}\mathbf{v}_{21}, \tag{15}$$

$$\mathbf{x}_2 = d_{12}\mathbf{v}_{12} + d_{22}\mathbf{v}_{22}, \tag{16}$$

where the unit vectors $\mathbf{v}_{rt} \in \mathcal{C}^{3\times 1}$, $r,t=1,2$, denote the modulation vectors. The power of p_{rt} is allocated to the data stream d_{rt}.

As mentioned in the previous section, the interference of d_{11} over d_{21}, and the interference of d_{22} over d_{12} are canceled out based on the dirty-paper-coding theorem. Motivated by the proof of the dirty-paper-coding theorem in [6], we embed data in $\hat{d}_{21}$ and $\hat{d}_{12}$, where

$$\begin{aligned}
d_{21} &= \hat{d}_{21} - \alpha_2(\mathbf{v}_{21}^{\dagger}\overline{\mathbf{H}}_{21}^{\dagger}\overline{\mathbf{H}}_{21}\mathbf{v}_{21})^{-1}\mathbf{v}_{21}^{\dagger}\overline{\mathbf{H}}_{21}^{\dagger}\overline{\mathbf{H}}_{21}\mathbf{v}_{11}d_{11}\\
d_{12} &= \hat{d}_{12} - \alpha_1(\mathbf{v}_{12}^{\dagger}\overline{\mathbf{H}}_{12}^{\dagger}\overline{\mathbf{H}}_{12}\mathbf{v}_{12})^{-1}\mathbf{v}_{12}^{\dagger}\overline{\mathbf{H}}_{12}^{\dagger}\overline{\mathbf{H}}_{12}\mathbf{v}_{22}d_{22}\\
\alpha_1 &= p_{12}\left(p_{12} + (\mathbf{v}_{12}^{\dagger}\overline{\mathbf{H}}_{12}^{\dagger}\overline{\mathbf{H}}_{12}\mathbf{v}_{12})^{-1}\right)^{-1},\\
\alpha_2 &= p_{21}\left(p_{21} + (\mathbf{v}_{21}^{\dagger}\overline{\mathbf{H}}_{21}^{\dagger}\overline{\mathbf{H}}_{21}\mathbf{v}_{21})^{-1}\right)^{-1},
\end{aligned}$$

where $(.)^{\dagger}$ denotes transpose conjugate operation. $\overline{\mathbf{H}}_{21}$ and $\overline{\mathbf{H}}_{12}$ are defined later in (20) and (26).

As mentioned, at the receiver side, the successive decoding (SD) scheme is employed. The structure of the receiver is as follows: first, user 1 decodes $\hat{d}_{12}$ and subtracts its effect from the received vector $\mathbf{y}_1$. Then, d_{11} is decoded. Similarly, user 2 first decodes $\hat{d}_{21}$ and subtracts its effect from $\mathbf{y}_2$, then decodes d_{22}. The details of the detection are depicted in Fig. 2. To decode $\hat{d}_{12}$ at the user 1 terminal, the signals received from base station 1 , i.e. d_{11} and d_{21}, are treated as interference. The proposed precoding scheme is such that the data stream d_{22} has no interference on the data stream $\hat{d}_{12}$. The filter $\mathbf{\Psi}_{12} = \mathbf{R}_{12}^{-\frac{1}{2}}$ is used to whiten the interference plus noise $\mathbf{H}_{11}(\mathbf{v}_{11}d_{11} + \mathbf{v}_{21}d_{21}) + \mathbf{n}_1$ with the variance matrix $\mathbf{R}_{12}$,

$$\mathbf{R}_{12} = \mathbf{H}_{11}[\mathbf{v}_{11}\ \mathbf{v}_{21}]\begin{bmatrix} p_{11} & 0\\ 0 & p_{21}\end{bmatrix}[\mathbf{v}_{11}\ \mathbf{v}_{21}]^{\dagger}\mathbf{H}_{11}^{\dagger} + \mathbf{I}. \tag{17}$$

This formula is based on the result in [6] which implies d_{11} and d_{21} are independent.

The output of $\mathbf{\Psi}_{12}$ is passed through the filter $\mathbf{u}_{12}$ which maximizes the effective SINR. The design of the precoding and the filter $\mathbf{u}_{12}$ will be explained later. Here, the user one decodes $\hat{d}_{12}$ and then subtracts its effect from the received signal $\mathbf{y}_1$, i.e.

$$\begin{aligned}
\widetilde{\mathbf{y}}_1 &= \mathbf{y}_1 - \mathbf{H}_{12}\mathbf{v}_{12}\hat{d}_{12}\\
&= \mathbf{Q}_1\mathbf{H}_{12}\mathbf{v}_{22}d_{22} + \mathbf{H}_{11}\mathbf{v}_{11}d_{11} + \mathbf{H}_{12}\mathbf{v}_{12}d_{12}
\end{aligned}$$

where,

$$\mathbf{Q}_1 = \mathbf{I} - \mathbf{H}_{12}\mathbf{v}_{12}\alpha_1(\mathbf{v}_{12}^{\dagger}\overline{\mathbf{H}}_{12}^{\dagger}\overline{\mathbf{H}}_{12}\mathbf{v}_{12})^{-1}\mathbf{v}_{12}^{\dagger}\overline{\mathbf{H}}_{12}^{\dagger}\mathbf{\Psi}_{12}$$

In the next step, the user one decodes d_{11} from $\widetilde{\mathbf{y}}_1$. First, the filter $\mathbf{\Psi}_{11}$ is used to whiten the interference of d_{22} over d_{11}. Note that the data stream d_{21} has no interference over d_{11} due to the precoding at the transmitter. The interference plus noise is equal to $\mathbf{Q}_1\mathbf{H}_{12}\mathbf{v}_{22}d_{22} + \mathbf{n}_1$ with the covariance matrix $\mathbf{R}_{11} = \mathbf{Q}_1\mathbf{H}_{12}\mathbf{v}_{22}p_{22}\mathbf{v}_{22}^{\dagger}\mathbf{H}_{12}^{\dagger}\mathbf{Q}_1^{\dagger} + \mathbf{I}$. Then, the whitening filter is equal to $\mathbf{\Psi}_{11} = \mathbf{R}_{11}^{-\frac{1}{2}}$. The output of the whitening filter $\mathbf{\Psi}_{11}$ is passed through the filter $\mathbf{u}_{11}^{\dagger}$ which maximizes the SNR of the data stream d_{11}. Similarly, for the user 2, there are two whitening filters

$\mathbf{\Psi}_{21} = \mathbf{R}_{21}^{-\frac{1}{2}}$ and $\mathbf{\Psi}_{22} = \mathbf{R}_{21}^{-\frac{1}{2}}$ where,

$$\mathbf{R}_{21} = \mathbf{H}_{22}[\mathbf{v}_{12}\ \mathbf{v}_{22}] \begin{bmatrix} p_{12} & 0 \\ 0 & p_{22} \end{bmatrix} [\mathbf{v}_{12}\ \mathbf{v}_{22}]^\dagger \mathbf{H}_{22}^\dagger + \mathbf{I}$$
$$\mathbf{R}_{22} = \mathbf{Q}_2 \mathbf{H}_{21} \mathbf{v}_{11} p_{11} \mathbf{v}_{11}^\dagger \mathbf{H}_{21}^\dagger \mathbf{Q}_2^\dagger + \mathbf{I},$$
$$\mathbf{Q}_2 = \mathbf{I} - \mathbf{H}_{21}\mathbf{v}_{21}\alpha_2(\mathbf{v}_{21}^\dagger \overline{\mathbf{H}}_{21}^\dagger \overline{\mathbf{H}}_{21}\mathbf{v}_{21})^{-1}\mathbf{v}_{21}^\dagger \overline{\mathbf{H}}_{21}^\dagger \mathbf{\Psi}_{21}$$

Similarly, at the user two terminal, $\mathbf{u}_{21}^\dagger$ and $\mathbf{u}_{22}^\dagger$ are used to detect d_{21} and d_{22}, respectively. Figure 2 shows some more details.

In what follows, we explain the derivation of the modulation vectors $\mathbf{v}_{rt}$ and the demodulation vectors $\mathbf{u}_{rt}$, $r, t = 1, 2$. To this end, we consider the second perspective of the system as a set of two broadcast channels. As depicted in Fig. 2, the following MIMO broadcast channel is viewed from the base station one,

$$\widehat{\mathbf{y}}_1 = \overline{\mathbf{H}}_{11}\mathbf{x}_1 + \widehat{\mathbf{n}}_1, \tag{18}$$
$$\check{\mathbf{y}}_2 = \overline{\mathbf{H}}_{21}\mathbf{x}_1 + \check{\mathbf{n}}_2, \tag{19}$$

where $\widehat{\mathbf{n}}_1$ and $\check{\mathbf{n}}_2$ are whitened noise terms and

$$\overline{\mathbf{H}}_{11} = \mathbf{\Psi}_{11}\mathbf{H}_{11}, \tag{20}$$
$$\overline{\mathbf{H}}_{21} = \mathbf{\Psi}_{21}\mathbf{H}_{21}. \tag{21}$$

For signaling, we apply the scheme proposed in [14] for the MIMO broadcast systems with multiple receive antennas. According to [14], the modulation vector $\mathbf{v}_{11}$ is equal to the optimizing vector of the following maximization problem,

$$\begin{aligned} \sigma_{11}^2 = \quad & \max_{\mathbf{s}} \mathbf{s}^\dagger \overline{\mathbf{H}}_{11}^\dagger \overline{\mathbf{H}}_{11}\mathbf{s}, \\ s.t. \quad & \mathbf{s}^\dagger\mathbf{s} = 1 \end{aligned} \tag{22}$$

where σ_{11} is the gain of the equivalent single-antenna channel on which the data stream d_{11} is sent. The demodulation vector $\mathbf{u}_{11}$ is given by $\mathbf{u}_{11} = \frac{\overline{\mathbf{H}}_{11}\mathbf{v}_{11}}{\sigma_{11}}$.

$\mathbf{v}_{21}$ is the optimizing vector of the following maximization problem,

$$\begin{aligned} \sigma_{21}^2 = \quad & \max_{\mathbf{s}} \mathbf{s}^\dagger \overline{\mathbf{H}}_{21}^\dagger \overline{\mathbf{H}}_{21}\mathbf{s}, \\ s.t. \quad & \mathbf{s}^\dagger\mathbf{s} = 1, \quad \mathbf{v}_{11}^\dagger\mathbf{s} = 0 \end{aligned} \tag{23}$$

where σ_{21} is the gain of the equivalent channel on which the data stream d_{21} is sent. The demodulation vector $\mathbf{u}_{21}$ is given by $\mathbf{u}_{21} = \frac{\overline{\mathbf{H}}_{21}\mathbf{v}_{21}}{\sigma_{21}}$. As shown in [14], by using this scheme, the data stream d_{21} has no interference over the data stream d_{11}. As mentioned, knowing the selected codeword for data stream d_{11}, the base station one can effectively cancel out the interference of the data stream d_{11} over d_{21} based on the dirty-paper coding theorem. Consequently, the broadcast channel is reduced to a set of two parallel channels with gains σ_{11} and σ_{21}. Water-filling is applied to optimally allocate the powers p_{11} and p_{21} to the data streams d_{11} and d_{21}, respectively, where $p_{11} + p_{21} \leq P$.

Similarly, from the base station 2, we have a MIMO broadcast channel modeled by

$$\check{\mathbf{y}}_1 = \overline{\mathbf{H}}_{12}\mathbf{x}_2 + \check{\mathbf{n}}_1, \tag{24}$$
$$\widehat{\mathbf{y}}_2 = \overline{\mathbf{H}}_{22}\mathbf{x}_2 + \widehat{\mathbf{n}}_2, \tag{25}$$

where $\check{\mathbf{n}}_1$ and $\widehat{\mathbf{n}}_2$ are whitened noises and

$$\overline{\mathbf{H}}_{12} = \mathbf{\Psi}_{12}\mathbf{H}_{12} \tag{26}$$
$$\overline{\mathbf{H}}_{22} = \mathbf{\Psi}_{22}\mathbf{H}_{22}. \tag{27}$$

Here, we apply the same algorithm to derive the modulation and demodulation vectors for the base station 2. $\mathbf{v}_{22}$ is equal to the optimizing vector of the following maximization problem

$$\begin{aligned} \sigma_{22}^2 = \quad & \max_{\mathbf{s}} \mathbf{s}^\dagger \overline{\mathbf{H}}_{22}^\dagger \overline{\mathbf{H}}_{22}\mathbf{s}, \\ s.t. \quad & \mathbf{s}^\dagger\mathbf{s} = 1 \end{aligned}$$

and $\mathbf{u}_{22} = \frac{\overline{\mathbf{H}}_{22}\mathbf{v}_{22}}{\sigma_{22}}$. In addition, $\mathbf{v}_{12}$ is equal to the optimizing vector in the following problem

$$\begin{aligned} \sigma_{12}^2 = \quad & \max_{\mathbf{s}} \mathbf{s}^\dagger \overline{\mathbf{H}}_{12}^\dagger \overline{\mathbf{H}}_{12}\mathbf{s}, \\ s.t. \quad & \mathbf{s}^\dagger\mathbf{s} = 1, \quad \mathbf{v}_{22}^\dagger\mathbf{s} = 0 \end{aligned}$$

and $\mathbf{u}_{12} = \frac{\overline{\mathbf{H}}_{12}\mathbf{v}_{12}}{\sigma_{12}}$. Similar to the first base station, d_{12} has no interference over d_{22}. Selecting the codeword for d_{22}, the transmitter two can effectively cancel its interference over d_{12}, using the dirty paper coding. Water-filling is used to optimally divide the total power P between p_{12} and p_{22}. At the end, this method reduces the system to four parallel channels with the channel gains σ_{rt}, $r, t = 1, 2$. Therefore, the sum-rate of the proposed scheme is obtained by,

$$R_{Sum-Rate} = \sum_{r=1}^{2}\sum_{t=1}^{2} \log_2(1 + \sigma_{rt}^2 p_{rt}). \tag{28}$$

Note that to compute $\mathbf{v}_{11}$ and $\mathbf{v}_{12}$ in (22) and (23), $\mathbf{v}_{21}$ and $\mathbf{v}_{22}$ are needed ($\mathbf{\Psi}_{11}$ and $\mathbf{\Psi}_{21}$ are functions of $\mathbf{v}_{21}$ and $\mathbf{v}_{22}$) and vise versa. To derive the modulation vectors, we can initialize $\mathbf{v}_{rt}$, $r, t = 1, 2$, randomly, and iteratively follow (22) to (28), until the resulting vectors converge. Simulation results show that the algorithm converges very fast.

IV. Performance Analysis

Although finding the optimal power allocation is straight-forward, to simplify the analysis, we assume that each base station divides the total power equally between data streams, i.e. $p_{rt} = P/2$, $t, r = 1, 2$. In this case, we have, $\mathbf{R}_{11} = \frac{P}{2}\mathbf{Q}_1\mathbf{H}_{12}\mathbf{v}_{22}\mathbf{v}_{22}^\dagger\mathbf{H}_{12}^\dagger\mathbf{Q}_1^\dagger + \mathbf{I}$. Let $\delta_1 = |\mathbf{Q}_1\mathbf{H}_{12}\mathbf{v}_{22}|$ and $\overline{\mathbf{v}}_{22} = \frac{\mathbf{Q}_1\mathbf{H}_{12}\mathbf{v}_{22}}{|\mathbf{Q}_1\mathbf{H}_{12}\mathbf{v}_{22}|}$. Consider the unit vectors $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ such that $[\overline{\mathbf{v}}_{22}, \boldsymbol{\nu}_1, \boldsymbol{\nu}_2]$ forms a unitary matrix. Then, we can show that,

$$\overline{\mathbf{H}}_{11} = \begin{bmatrix} \frac{1}{\sqrt{\frac{P}{2}\delta_1^2+1}} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \overline{\mathbf{v}}_{22}^\dagger \\ \boldsymbol{\nu}_1^\dagger \\ \boldsymbol{\nu}_2^\dagger \end{bmatrix} \mathbf{H}_{11}. \tag{29}$$

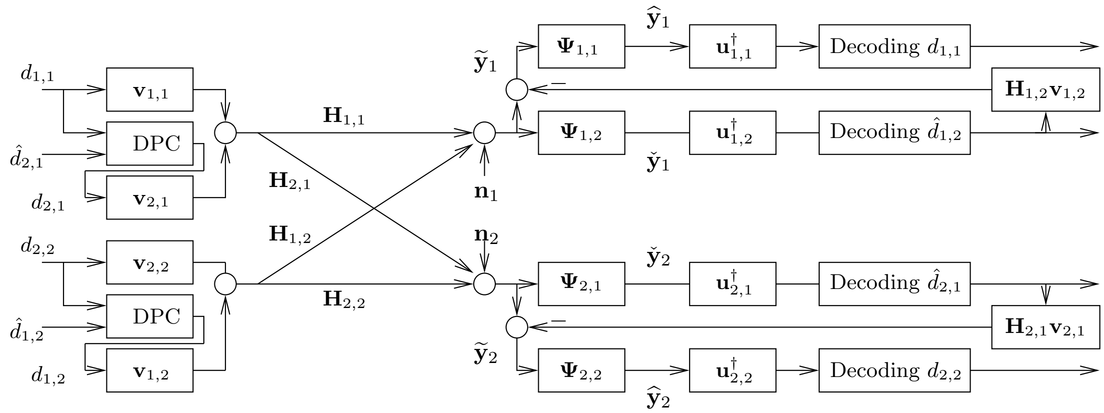

Fig. 2. Block Diagram of the Proposed Precoding and Detection Schemes

In high SNR, $\frac{1}{\sqrt{\frac{P}{2}\delta_1^2+1}} \rightarrow 0$. Thus, we have $\overline{\mathbf{H}}_{11} = [\mathbf{0}\ \boldsymbol{\nu}_1\ \boldsymbol{\nu}_2]^{\dagger}\mathbf{H}_{11}$. Regarding (22), σ_{11} is equal to the maximum singular value of $\overline{\mathbf{H}}_{11}$, which is a rank 2 matrix for large SNR. Therefore, σ_{11} converges to a non-vanishing positive constant. Similar statements are valid for σ_{22}. Consequently, each of the data streams d_{11} and d_{22} achieves multiplexing gain of one. Now, we investigate the multiplexing gain of the data streams d_{12} and d_{21}. Let $p_{11} = p_{21} = \frac{P}{2}$. From (17), we have

$$\mathbf{R}_{12} = \frac{P}{2}\mathbf{H}_{11}[\mathbf{v}_{11}\ \mathbf{v}_{21}][\mathbf{v}_{11}\ \mathbf{v}_{21}]^{\dagger}\mathbf{H}_{11}^{\dagger} + \mathbf{I}.$$

Applying the SVD decomposition, we have $\mathbf{H}_{11}[\mathbf{v}_{11}\ \ \mathbf{v}_{21}][\mathbf{v}_{11}\ \ \mathbf{v}_{21}]^{\dagger}\mathbf{H}_{11}^{\dagger} = \lambda_1^2\boldsymbol{\varpi}_1\boldsymbol{\varpi}_1^{\dagger} + \lambda_2^2\boldsymbol{\varpi}_2\boldsymbol{\varpi}_2^{\dagger}$, where $\lambda_1, \lambda_2 \geq 0$, and $\boldsymbol{\varpi}_1$ and $\boldsymbol{\varpi}_2$ are two unit orthogonal vectors. Consider $\boldsymbol{\varpi}_3$ such that the matrix $[\boldsymbol{\varpi}_1, \boldsymbol{\varpi}_2, \boldsymbol{\varpi}_3]$ forms a unitary matrix, then we can show that

$$\overline{\mathbf{H}}_{12} = \begin{bmatrix} \frac{1}{\sqrt{\frac{P}{2}\lambda_1^2+1}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{\frac{P}{2}\lambda_2^2+1}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{\varpi}_1^{\dagger} \\ \boldsymbol{\varpi}_2^{\dagger} \\ \boldsymbol{\varpi}_3^{\dagger} \end{bmatrix} \mathbf{H}_{12}. \quad (30)$$

As it is shown in [14], σ_{12} in (23) is equal to the maximum singular value of $\overline{\overline{\mathbf{H}}}_{12}$, where $\overline{\overline{\mathbf{H}}}_{12} = \overline{\mathbf{H}}_{12}[\boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2]$, and $\boldsymbol{\varphi}_1$ and $\boldsymbol{\varphi}_2$ are two unit vectors such that $[\mathbf{v}_{11}, \boldsymbol{\varphi}_1, \boldsymbol{\varphi}_2]$ forms a unitary matrix. In high SNR, $\frac{1}{\sqrt{\frac{P}{2}\lambda_1^2+1}} \rightarrow 0$ and $\frac{1}{\sqrt{\frac{P}{2}\lambda_2^2+1}} \rightarrow 0$. Consequently, $\overline{\overline{\mathbf{H}}}_{12}$ converges to a matrix with rank one. Therefore, σ_{21}, defined in (23), converges to non-vanishing positive number. Thus, the data stream d_{21} archives multiplexing gain of one. Similar statements are valid for d_{12}.

Theorem 2 *In a MIMO system with two transmitters and two receivers, each of them equipped with three antennas, the proposed scheme achieves multiplexing gain of four.*

As mentioned, if we apply conventional schemes for this system, the maximum achievable multiplexing gain is three [4]. The above result clearly shows the advantage of the proposed scheme. Note that in this scheme, the signals of the transmitter one and two are uncorrelated. In fact, the only information which has to be shared between the base stations are all the channel matrices.

References

[1] B. S. Blum, "MIMO capacity with interference," *IEEE Journal on Selected Areas in Communications*, vol. 21, pp. 793 – 801, June 2003.

[2] H. Dai, A. F. Molisch, and H. V. Poor, "Downlink capacity of interference-limited MIMO systems with joint detection," *IEEE Transactions on Wireless Communications*, vol. 3, pp. 442 – 453, March 2004.

[3] H. Dai and H. V. Poor, "Asymptotic spectral efficiency of multicell MIMO systems with frequency-flat fading," *IEEE Transactions on Signal Processing*, vol. 51, pp. 2976 – 2988, Nov. 2003.

[4] S. A. Jafar, "Degrees of freedom in distributed mimo communications," in *IEEE Communication Theory Workshop*, 2005.

[5] S. Shamai (Shitz) and B.M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *IEEE Vehicular Technology Conference*, May 2001, vol. 3, pp. 1745 – 1749.

[6] M. Costa, "Writing on dirty paper," *IEEE Trans. Inform. Theory*, vol. 29, pp. 439–441, May 1983.

[7] G. J. Foschini, H. Huang, K. Karakayali, R. A. Valenzuela, and S. Venkatesan, "The value of coherent base station coordination," in *Conference on Information Sciences and Systems (CISS)*, March 2005.

[8] S. A. Jafar and A. J. Goldsmith, "Transmitter optimization for multiple antenna cellular systems," in *IEEE International Symposium on Information Theory*, 2002, p. 50.

[9] C. Rose O. Popescu, "Sum capacity and TSC bounds in collaborative multi-base wireless systems," *IEEE Transactions on Information Theory*, vol. 10, pp. 2433–2438, October 2004.

[10] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Transactions on Information Theory*, vol. 25, pp. 306–311, May 1979.

[11] A. El Gamal and E. van der Meulen, "A proof of marton's coding theorem for the discrete memoryless broadcast channel (corresp.)," *IEEE Transactions on Information Theory*, vol. 27, pp. 120–122, Jan. 1981.

[12] H. Liao, *Multiple access channels*, Ph.D. thesis, Dep. Elec. Eng., Univ. of Hawaii., 1972.

[13] S. Gel'fand and M. Pinsker, "Coding for channel with random parameters," *Problems of Control and Information Theory*, vol. 9(1), pp. 19–31, 1980.

[14] M. A. Maddah-Ali, M. Ansari, and Amir K. Khandani, "An efficient signaling scheme for MIMO broadcast systems: Design and performance evaluation," *IEEE Transactions on Information Theory*, July 2005, Submitted for Publication.