# Optimal Antenna Selection Based on Capacity Maximization for MIMO Systems in Correlated Channels\*

Lin Dai, Member, IEEE, Sana Sfar, Student Member, IEEE, and Khaled B. Letaief, Fellow, IEEE

Center for Wireless Information Technology Electrical and Electronic Engineering Department The Hong Kong University of Science & Technology Clear Water Bay, HONG KONG

# Abstract

Recent work has shown that multiple-input multiple-output (MIMO) systems with multiple antennas at both the transmitter and receiver are able to achieve great capacity improvement. In such systems, it is desirable to select a subset of the available antennas so as to reduce the number of RF chains. This paper addresses the problem of optimal antenna selection in correlated channels. We consider a narrow-band communication system with M transmit and N receive antennas. We present the criterion for selecting the optimal  $L_t$  out of M transmit and  $L_r$  out of N receive antennas in terms of capacity maximization, assuming that only the long-term channel statistics, instead of the instantaneous channel state information, are known. Simulations will be used to validate our theoretical analysis and demonstrate that the number of required RF chains can be significantly decreased using our proposed selection strategy while achieving even better performance than the conventional MIMO system without antenna selection.

**Keywords:** Correlated channels, Antenna selection, MIMO systems, Channel capacity, Singular value decomposition.

**Corresponding Author:** 

Professor K. B. Letaief Email: <u>eekhaled@ee.ust.hk</u> Tel:/Fax: (852) 2358 7064/1485 URL: <u>http://www.ee.ust.hk/~eekhaled</u>

<sup>\*</sup> This work is supported in part by the Hong Kong Telecom Institute of Information Technology.

## I. Introduction

MIMO (Multiple-Input Multiple-Output) systems have recently attracted tremendous interest due to their ability in providing great capacity improvements [1-2]. In particular, a new technology denoted by layered space-time for MIMO systems (BLAST) has been proposed in [3]. Such scheme is shown to achieve unprecedented capacities that grow linearly with the number of transmit and receive antennas, when all signals undergo independent fading. However, the deployment of multiple antennas would require the implementation of multiple RF chains that are typically very expensive. Dealing with this issue, [5] firstly proposed to select only the most useful antennas for further signal processing, namely, only L out of N antennas are effectively deployed, and only L RF chains are thus required. In [5-6], a system known as Hybrid Selection/Maximum ratio combining is proposed in which the antennas are selected to maximize the achieved diversity gain as well as minimize the obtained error rates. However, all of these previous studies are based on the multiple-input single-output (MISO) channel or the single-input multiple-output (SIMO) channel. [7-10] further applied the antenna selection to MIMO links and showed that in a multiple-antenna fading channel, antenna selection can also provide diversity advantage. Various criteria for receive antenna selection or transmit antenna selection were proposed aiming at minimizing the symbol error rate [7] or maximizing the capacity bounds [8-10].

The above stated results hold unfortunately only when the channel is rich enough. In such case, the transmitted data is split into several streams and transmitted in parallel over individual and independent channel links so that spatial multiplexing gain can be obtained. In fact, such assumption is generally not realistic and channel links usually present spatial correlation<sup>1</sup> due to the lack of spacing between antennas, or to the existence of small angular spread. Both cases lead to a diminishing diversity and multiplexing gain, and this will significantly affect the capacity and error probability performance [11]. Particularly, when BLAST is applied in correlated channels, the performance is severely degraded to an unacceptable level since there are not enough independent dimensions supporting the simultaneously transmitted streams. Therefore, some processing at the transmitter must be done to combat this harmful effect. In this paper, we will show that in the correlated scenario, proper transmit antenna selection can not only be used to decrease the number of

<sup>&</sup>lt;sup>1</sup> Throughout this paper, the term "correlation" shall refer to "spatial correlation".

RF chains, but also as an effective means to improve the performance.

We consider a narrow-band communication system with M transmit and N receive antennas over a slowly varying flat Rayleigh fading correlated channel. We propose to select only  $L_r$  out of M transmit and  $L_r$  out of N receive antennas for further signal processing. We present the optimal antenna selection criterion for capacity maximization assuming that only the long-term channel statistics (LT-CS) are available. These statistics only change with the antenna position patterns or the surrounding environment, and thus remain invariant for long time intervals. The proposed selection process is not, hence, updated for each channel instance as those presented in [7-10], which are all based on the exact instantaneous channel state information (IC-SI). Moreover, since LT-CS are given by the correlation matrices at both ends, no feedback channel is necessary for transmit antenna selection. Therefore, our proposed algorithm, which we shall refer to as *correlated selection algorithm* (CSA) has the advantage of introducing further complexity reduction than that using IC-SI which we refer to as *instantaneous selection algorithm* (ISA).

Unlike the work in [13-15] which focus on the selection criterion for minimizing the average error probability, here we demonstrate that for capacity maximization, the transmit (receive) antenna subset should be chosen to maximize the determinant of the transmit (receive) correlation matrix. To do so, we derive the capacity upper and lower bounds, and show that they converge to the same limit. By maximizing both bounds, we maximize the capacity and obtain the optimal selection criterion. Simulation results are used to further confirm our analysis and show that with our proposed CSA, significant gain can be achieved over the random selection scheme, especially at the transmitter side. Besides, comparison with ISA will indicate only slight capacity degradation. This implies that in correlated channels, antenna selection can be based on LT-CS instead of IC-SI while keeping capacity levels nearly unchanged. We also consider the case of the conventional system<sup>2</sup> and find that CSA can even achieve better performance thanks to its optimal transmit selection.

As for the joint antenna selection at both communication ends, we show that such a process can be decoupled with CSA and significant complexity reduction is consequently obtained at satisfying levels. For a large M or N, however, CSA still involves high computational complexity levels due to the required exhaustive search for the global optimal antenna set. To deal with this issue, we propose a

 $<sup>^2</sup>$  Throughout the paper, the "conventional system" shall refer to the system that uses all the *M* transmit and *N* receive antennas without selecting any antenna subset.

new low-complexity selection algorithm, denoted by L-CSA. It consists of iteratively searching the local optimal antenna subset at each stage. Significant complexity reduction as well as very close performance to CSA, are shown to be achieved.

This paper is organized as follows. In Section II, we provide the system model and the various notations used throughout the paper. In Section III, we derive the antenna selection criterion for capacity maximization based on LT-CS. Section IV presents the details of our selection algorithm, CSA, and the complexity analysis. Another low-complexity antenna selection algorithm, L-CSA, is also proposed. Section V shows the capacity performance of both CSA and L-CSA. Comparison results with ISA, as well as, the conventional system are also provided in this section. Finally, Section VI summarizes and concludes this paper.

## **II. System Model**

We consider in this paper the transmission of  $L_t$  signals through M antennas ( $M \ge L_t$ ), which undergo a slowly Rayleigh fading correlated channel to reach a receiver with N antennas. The received signals are next multiplexed into  $L_r$  ( $L_r \le N$ ) RF chains so as to reduce the receiver cost and complexity. For simplicity, we assume in the following a perfect channel knowledge at the receiver side only, through the use of training sequences. We also assume that the channel is constant within a frame of T symbols.

Let **H** denote the  $N \times M$  channel matrix and  $\mathbf{x} = [x_1, x_2, ..., x_M]^{'}$  denote the transmitted signal vector, where  $x_i$  is the transmitted symbol from the  $i^{th}$  antenna and  $(\cdot)^{'}$  refers to the transpose operator. Assuming perfect symbol synchronization at the receiver as well as equal transmission power at the transmitter side, the discrete model of the received complex signal vector can be written as

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1}$$

where **n** denotes a complex Gaussian *N*-vector noise with covariance  $\sigma^2 \mathbf{I}_{N \times N}$ .

Following the channel model provided in [11,12], the channel matrix could be written as

$$\mathbf{H} = \mathbf{R}_{r}^{1/2} \mathbf{H}_{w} \mathbf{R}_{t}^{1/2} \tag{2}$$

where  $\mathbf{H}_{w}$  is an  $N \times M$  complex matrix of i.i.d. zero-mean, unit variance complex Gaussian entries.  $\mathbf{R}_{t}$  and  $\mathbf{R}_{r}$  denote the  $M \times M$  and  $N \times N$  antenna correlation matrices at the transmitter and receiver side, respectively. It is also noted that these matrices have unit diagonal entries.

We define the selected transmit antenna subset and selected receive antenna subset as  $\Lambda_t$  and  $\Lambda_r$ , respectively, which are both unordered sets with  $L_t$  and  $L_r$  selected antennas. Let  $\tilde{\mathbf{y}}$ ,  $\tilde{\mathbf{x}}$ , and  $\tilde{\mathbf{n}}$  be the receive signal, transmit signal and noise vector after selection, respectively. Let also  $\mathbf{R}_{\Lambda_r}$  and  $\mathbf{R}_{\Lambda_t}$  denote the cross-correlation matrix of those  $L_r$  and  $L_t$  selected antennas, respectively. These matrices can be obtained by eliminating the columns and rows of the non-desired antennas from  $\mathbf{R}_r$  and  $\mathbf{R}_t$ , respectively. We assume here that  $L_t$  and  $L_r$  are selected to satisfy that  $\mathbf{R}_{\Lambda_t}$  and  $\mathbf{R}_{\Lambda_r}$  are both full ranks. Finally let  $\tilde{\mathbf{H}}$  represent the  $L_r \times L_t$  channel gain matrix between  $L_t$  selected transmit and  $L_r$  selected receive antennas. Then,

$$\tilde{\mathbf{y}} = \tilde{\mathbf{H}}\tilde{\mathbf{x}} + \tilde{\mathbf{n}} = \mathbf{R}_{\Lambda_{r}}^{1/2}\tilde{\mathbf{H}}_{w}\mathbf{R}_{\Lambda_{t}}^{1/2}\tilde{\mathbf{x}} + \tilde{\mathbf{n}} .$$
(3)

Throughout this paper, we denote by  $(\cdot)^*$ , det $(\cdot)$ , trace $(\cdot)$  and rank $(\cdot)$  the complex conjugate transpose, the determinant, the trace and the rank operators, respectively.  $\mathbf{0}_{men}$  represents an  $m \times n$  zero matrix. For an arbitrary matrix  $\mathbf{A}$ ,  $a_{ij}$  refers to its element at the  $i^{th}$  row and the  $j^{th}$  column,  $\lambda_{\mathbf{A}}^{(i)}$  to its  $i^{th}$  eigenvalue<sup>3</sup>, and  $(\mathbf{A})_{n\times n}$  to its  $n \times n$  principal submatrix. When  $\mathbf{A}$  is a Hermitian non-negative definite matrix, we shall write it as  $\mathbf{A} \ge \mathbf{0}$ . We also write  $\mathbf{A}$  as  $diag(a_{ii})$  when it is a diagonal matrix. Finally, a set  $\mathbf{S}$  with elements  $s_i$ , i=1,2,...,n, will be represented by  $\{s_1,s_2,...,s_n\}$  with its length denoted by  $|\mathbf{S}|$ .

## **III. Antenna Selection Criterion**

In the following derivation, we highlight the effect of correlation on the capacity, so as to obtain a selection criterion that is related only to  $\mathbf{R}_{\Lambda_{r}}$  and  $\mathbf{R}_{\Lambda_{r}}$ . To do so, we start by applying *Singular* 

<sup>&</sup>lt;sup>3</sup> The eigenvalues are all assumed to be sorted in a descending order throughout this paper.

*Value Decomposition* (SVD) to  $\mathbf{R}_{\Lambda_r}$  and  $\mathbf{R}_{\Lambda_t}$ . We then obtain

$$\mathbf{R}_{\Lambda_r} = \mathbf{U}_r \mathbf{Q}_r \mathbf{U}_r^*, \quad \mathbf{R}_{\Lambda_t} = \mathbf{U}_t \mathbf{Q}_t \mathbf{U}_t^* \tag{4}$$

where  $\mathbf{U}_{r}$  and  $\mathbf{U}_{t}$  are both unitary matrices whose columns are the eigenvectors of  $\mathbf{R}_{\Lambda_{r}}$  and  $\mathbf{R}_{\Lambda_{t}}$ , respectively.  $\mathbf{Q}_{r}$  and  $\mathbf{Q}_{t}$  are both diagonal matrices whose diagonal entries are eigenvalues of  $\mathbf{R}_{\Lambda_{r}}$ and  $\mathbf{R}_{\Lambda_{t}}$ , respectively. The channel matrix,  $\tilde{\mathbf{H}}$ , is thus re-written as

$$\tilde{\mathbf{H}} = \mathbf{R}_{\Lambda_{\mathrm{r}}}^{1/2} \tilde{\mathbf{H}}_{\mathrm{w}} \mathbf{R}_{\Lambda_{\mathrm{t}}}^{1/2} = \mathbf{U}_{\mathrm{r}} \mathbf{Q}_{\mathrm{r}}^{1/2} \tilde{\mathbf{H}}_{\mathrm{w}} \mathbf{Q}_{\mathrm{t}}^{1/2} \mathbf{U}_{\mathrm{t}}^{*} \,.$$
(5)

From [4], we know that the capacity satisfies

$$C = \log_2 \det[\mathbf{I}_{\mathbf{L}_r \times \mathbf{L}_r} + \frac{\rho}{L_t} \tilde{\mathbf{H}} \tilde{\mathbf{H}}^*] = \log_2 \det[\mathbf{I}_{\mathbf{L}_r \times \mathbf{L}_r} + \frac{\rho}{L_t} \mathbf{U}_r \mathbf{Q}_r^{1/2} \tilde{\mathbf{H}}_w \mathbf{Q}_t \tilde{\mathbf{H}}^*_w \mathbf{Q}_r^{1/2} \mathbf{U}_r^*]$$

$$= \log_2 \det[\mathbf{I}_{\mathbf{L}_r \times \mathbf{L}_r} + \frac{\rho}{L_t} \tilde{\mathbf{H}}_w \mathbf{Q}_t \tilde{\mathbf{H}}^*_w \mathbf{Q}_r]$$
(6)

where  $\rho$  is the mean SNR per receive branch.

In order to select the optimal set of antennas that maximizes the above capacity expression, we distinguish three cases:  $L_r = L_t$ ,  $L_r > L_t$  and  $L_r < L_t$ . Actually, the  $L_r < L_t$  case is similar to  $L_r > L_t$  since  $\mathbf{Q}_r$  and  $\mathbf{Q}_t$  can be swapped with no effect on capacity as (6) shows. Therefore, in the following we will focus on the first two cases.

In the first case where  $L_r = L_t$ , the capacity is found to be equivalent to

$$C \approx \log_2 \det\left[\frac{\rho}{L_t} \tilde{\mathbf{H}}_{\mathbf{w}} \mathbf{Q}_t \tilde{\mathbf{H}}_{\mathbf{w}}^* \mathbf{Q}_r\right]$$

$$= L_t \log_2\left(\frac{\rho}{L_t}\right) + \log_2 \det[\tilde{\mathbf{H}}_{\mathbf{w}} \tilde{\mathbf{H}}_{\mathbf{w}}^*] + \log_2 \det[\mathbf{Q}_t] + \log_2 \det[\mathbf{Q}_r]$$
(7)

at high values of  $\rho$ . Consequently, it is clear that to maximize the capacity, we should maximize the determinants of  $\mathbf{R}_{\Lambda_t}$  and  $\mathbf{R}_{\Lambda_r}$ . In other words, the optimal transmit (receive) antenna set  $\Lambda_t$  ( $\Lambda_r$ ) in terms of capacity maximization should be selected to maximize the determinant of the corresponding correlation matrix  $\mathbf{R}_{\Lambda_t}$  ( $\mathbf{R}_{\Lambda_r}$ ). In the second scenario, however, it is difficult to obtain a closed form of the exact capacity expression. Thus, we propose to derive both a lower and an upper bound for the capacity so as to obtain the optimal selection criterion.

#### A. Capacity Lower Bound

From (6), we have

$$C = \log_2 \det[\mathbf{I}_{\mathbf{L}_t \times \mathbf{L}_t} + \frac{\rho}{L_t} \mathbf{Q}_t \tilde{\mathbf{H}}_w^* \mathbf{Q}_r \tilde{\mathbf{H}}_w]$$
(8)

given that  $\det(\mathbf{I} + \mathbf{AB}) = \det(\mathbf{I} + \mathbf{BA})$ . Let  $\mathbf{U}_{\mathbf{H}}$  denote an  $L_r \times L_r$  unitary matrix whose columns are the eigenvectors of  $\tilde{\mathbf{H}}_{\mathbf{w}}^* \tilde{\mathbf{H}}_{\mathbf{w}}$ , and  $\hat{\mathbf{Q}}_r$  the  $L_t \times L_t$  principal submatrix of  $\mathbf{U}_{\mathbf{H}}^* \mathbf{Q}_r \mathbf{U}_{\mathbf{H}}$ . It follows that

$$C > \log_{2} \det[\frac{\rho}{L_{t}} \mathbf{Q}_{t} \tilde{\mathbf{H}}_{w}^{*} \mathbf{Q}_{r} \tilde{\mathbf{H}}_{w}]$$

$$= L_{t} \log_{2} \left(\frac{\rho}{L_{t}}\right) + \log_{2} \det[\mathbf{Q}_{t}] + \log_{2} \det[\tilde{\mathbf{H}}_{w}^{*} \mathbf{Q}_{r} \tilde{\mathbf{H}}_{w}]$$

$$= L_{t} \log_{2} \left(\frac{\rho}{L_{t}}\right) + \log_{2} \det[\mathbf{Q}_{t}] + \log_{2} \det[\tilde{\mathbf{H}}_{w}^{*} \tilde{\mathbf{H}}_{w}] + \log_{2} \det[\hat{\mathbf{Q}}_{r}] \qquad (9)$$

Next, let  $\{q_r^{(j)}\}_{j \in \{1..L_r\}}$  denote the sorted eigenvalues in descending order of  $\mathbf{R}_{\Lambda_r}$ . Since  $\mathbf{U}_{\mathbf{H}}^* \mathbf{Q}_{\mathbf{r}} \mathbf{U}_{\mathbf{H}}$  is Hermitian, from [18], we know that  $\lambda_{\hat{\mathbf{Q}}_r}^{(i)} \ge \lambda_{\mathbf{U}_{\mathbf{H}}^* \mathbf{Q}_r \mathbf{U}_{\mathbf{H}}}^{(L_r - L_r + i)}$ ,  $i = 1, ..., L_t$ . As a result, we have

$$\log_2 \det[\hat{\mathbf{Q}}_{\mathbf{r}}] > \log_2 \prod_{i=L_r-L_r+1}^{L_r} q_r^{(i)} .$$
 (10)

By substituting (10) into (9), we have

$$\log_2 \det[\mathbf{Q}_t] + \log_2 \det[\tilde{\mathbf{H}}_w^* \tilde{\mathbf{H}}_w] + \log_2 \det[\hat{\mathbf{Q}}_r] > \log_2 \prod_{i=1}^{L_r} q_i^{(i)} + \log_2 \prod_{i=1}^{L_r} d_i + \log_2 \prod_{i=L_r-L_r+1}^{L_r} q_i^{(i)}, \quad (11)$$

where  $\{q_t^{(i)}\}_{i \in \{1, L_t\}}$  and  $\{d_i\}_{i \in \{1, L_t\}}$  are the sorted eigenvalues in descending order of  $\mathbf{R}_{\Lambda_t}$  and  $\tilde{\mathbf{H}}_{\mathbf{w}}^* \tilde{\mathbf{H}}_{\mathbf{w}}$ , respectively. Furthermore, it is easy to note that

$$\log_{2}\left(\prod_{i=L_{r}-L_{r}+1}^{L_{r}}q_{r}^{(i)}\right) = \log_{2}\left(\prod_{i=1}^{L_{r}}q_{r}^{(i)}\right) - \log_{2}\left(\prod_{i=1}^{L_{r}-L_{r}}q_{r}^{(i)}\right).$$
(12)

With the definition of  $\mathbf{R}_{\Lambda_r}$ , we know  $\sum_{i=1}^{L_r-L_r} q_r^{(i)} < L_r$  given that  $\sum_{i=1}^{L_r} q_r^{(i)} = L_r$ . Hence, by applying

Jensen's inequality, we obtain

$$\prod_{i=1}^{L_r - L_r} q_r^{(i)} \le \left(\frac{\sum_{i=1}^{L_r - L_r} q_r^{(i)}}{L_r - L_t}\right)^{L_r - L_r} < \left(\frac{L_r}{L_r - L_t}\right)^{L_r - L_r}.$$
(13)

Using (11), (12) and (13), we have

$$C = \log_2 \det[\mathbf{I}_{\mathbf{L}_t \times \mathbf{L}_t} + \frac{\rho}{L_t} \mathbf{Q}_t \tilde{\mathbf{H}}_w^* \mathbf{Q}_r \tilde{\mathbf{H}}_w]$$
  
>  $L_t \log_2\left(\frac{\rho}{L_t}\right) + \log_2\left(\prod_{i=1}^{L_t} d_i\right) + \log_2\left(\prod_{i=1}^{L_t} q_t^{(i)}\right) + \log_2\left(\prod_{i=1}^{L_r} q_r^{(i)}\right) - (L_r - L_t)\log_2\left(\frac{L_r}{L_r - L_t}\right).$  (14)

We thus obtain a lower bound,  $C_L$ , for the capacity that is given by

$$C_{L} = L_{t} \log_{2}\left(\frac{\rho}{L_{t}}\right) + (L_{r} - L_{t}) \log_{2}\left(\frac{L_{r} - L_{t}}{L_{r}}\right) + \log_{2}\left(\prod_{i=1}^{L_{t}} d_{i}\right) + \log_{2}\left(\prod_{i=1}^{L_{t}} q_{t}^{(i)}\right) + \log_{2}\left(\prod_{i=1}^{L_{r}} q_{r}^{(i)}\right).$$
(15)

#### **B.** Capacity Upper Bound

We begin by presenting the following lemma:

Lemma 1: Assume that **A** is an arbitrary  $n \times m$  matrix with non-negative eigenvalues and **B** is an  $m \times m$  non-negative definite diagonal matrix with  $m \le n$ . Then,  $\lambda_{ABA^*}^{(i)} \le \lambda_{A^*A}^{(1)} \lambda_B^{(i)}$ , i = 1, ..., m.

## Proof: See Appendix I.

Next, we note that

$$\det[\mathbf{I}_{\mathbf{L}_{t}\times\mathbf{L}_{t}} + \frac{\rho}{L_{t}}\mathbf{Q}_{t}\tilde{\mathbf{H}}_{w}^{*}\mathbf{Q}_{r}\tilde{\mathbf{H}}_{w}] = \det[\mathbf{I}_{\mathbf{L}_{r}\times\mathbf{L}_{r}} + \frac{\rho}{L_{t}}\mathbf{Q}_{r}\tilde{\mathbf{H}}_{w}\mathbf{Q}_{t}\tilde{\mathbf{H}}_{w}^{*}] = \det[\mathbf{I}_{\mathbf{L}_{r}\times\mathbf{L}_{r}} + \mathbf{Q}_{r}\mathbf{W}]$$
(16)

where  $\mathbf{W} = \frac{\rho}{L_t} \tilde{\mathbf{H}}_{\mathbf{w}} \mathbf{Q}_t \tilde{\mathbf{H}}_{\mathbf{w}}^*$ . Then, by applying the above lemma to  $\mathbf{W}$ , we have  $\lambda_{\mathbf{W}}^{(i)} \leq \frac{\rho}{L_t} \lambda_{\tilde{\mathbf{H}}_{\mathbf{w}}^* \tilde{\mathbf{H}}_{\mathbf{w}}}^{(i)} \lambda_{\mathbf{Q}_t}^{(i)}$  or

$$\lambda_{\mathbf{W}}^{(i)} \leq \frac{\rho}{L_{t}} d_{1} q_{t}^{(i)}, \quad i = 1, ..., L_{t}.$$

Now consider the  $L_r \times L_r$  diagonal matrix  $\mathbf{Z} = diag(z_{ii})$  with

$$z_{ii} = \begin{cases} \frac{\rho}{L_t} d_1 q_t^{(i)} & i = 1, ..L_t \\ 1 & i = L_t + 1, ..L_r \end{cases}$$
(17)

Clearly, we have  $z_{ii} \ge \lambda_{\mathbf{W}}^{(i)}$ ,  $\forall i = 1, ..., L_r$ . Then, we can write the following

$$\det[\mathbf{I}_{\mathbf{L}_{\mathbf{r}}\times\mathbf{L}_{\mathbf{r}}} + \mathbf{Q}_{\mathbf{r}}\mathbf{W}] \le \det[\mathbf{I}_{\mathbf{L}_{\mathbf{r}}\times\mathbf{L}_{\mathbf{r}}} + \mathbf{Q}_{\mathbf{r}}\mathbf{Z}] = \prod_{i=1}^{L_{\mathbf{r}}} (1 + q_{r}^{(i)}z_{ii}).$$
(18)

Finally, since  $\mathbf{R}_{\Lambda_r}$  is of full-rank, we have at high values of  $\rho$ ,

8

$$\prod_{i=1}^{L_r} (1+q_r^{(i)} z_{ii}) \approx \left(\frac{\rho}{L_t}\right)^{L_t} d_1^{L_t} \prod_{i=1}^{L_t} q_t^{(i)} \prod_{i=1}^{L_r} q_r^{(i)} .$$
(19)

The capacity is thus upper bounded by  $C_U$  where

$$C_{U} = L_{t} \log_{2}\left(\frac{\rho}{L_{t}}\right) + L_{t} \log_{2}\left(d_{1}\right) + \log_{2}\left(\prod_{i=1}^{L_{t}} q_{t}^{(i)}\right) + \log_{2}\left(\prod_{i=1}^{L_{r}} q_{r}^{(i)}\right).$$
(20)

#### C. Lower & Upper Bound Convergence

We will now show that the lower and upper bounds,  $C_L$  and  $C_U$ , converge to the same limit. To do so, we assume that  $L_t$  and  $L_r$  can increase without bound at the same rate, such that  $L_r \to \infty$ and  $L_t \to \infty$  with  $\frac{L_t}{L_r} \to r$ , as in [17]. Obviously r satisfies  $0 \le r < 1$ . We then show that  $\lim_{L_r \to \infty} \frac{1}{L_r} (C_U - C_L) \xrightarrow{r \to 0} 0$ .

Using the expressions of  $C_L$  and  $C_U$  given in (15) and (20), respectively, we have

$$\lim_{L_{r}\to\infty}\frac{1}{L_{r}}\left(C_{U}-C_{L}\right) = \lim_{L_{r}\to\infty}\left(\frac{L_{r}-L_{t}}{L_{r}}\log_{2}\frac{L_{r}}{L_{r}-L_{t}} + r\log_{2}d_{1} - \frac{1}{L_{r}}\sum_{i=1}^{L_{t}}\log_{2}d_{i}\right)$$
$$= \lim_{L_{r}\to\infty}\left(-(1-r)\log_{2}(1-r)\right) + \lim_{L_{r}\to\infty}\left(r\log_{2}d_{1} - \frac{1}{L_{r}}\sum_{i=1}^{L_{r}}\log_{2}d_{i}\right).$$
(21)

Now consider the limit of  $\left(r \log_2 d_1 - \frac{1}{L_r} \sum_{i=1}^{L_r} \log_2 d_i\right)$  in (21) when  $L_r \to \infty$ . To do so, let

 $\mathbf{S} = \frac{1}{L_t} \tilde{\mathbf{H}}_{\mathbf{w}} \tilde{\mathbf{H}}_{\mathbf{w}}^*, \text{ whose } i^{th} \text{ eigenvalue satisfies } \lambda_{\mathbf{S}}^{(i)} = \frac{1}{L_t} \lambda_{\tilde{\mathbf{H}}_{\mathbf{w}}^* \tilde{\mathbf{H}}_{\mathbf{w}}}^{(i)} = \frac{1}{L_t} d_i, \text{ for any } i = 1, ..., L_t. \text{ According to } i = 1, ..., L_t.$ 

[16], these eigenvalues satisfy

$$\lim_{L_r \to \infty} \max_{i \in \{1..L_r\}} \lambda_{\mathbf{S}}^{(i)} \to (1 + \sqrt{r})^2.$$
(22)

In addition, they have an empirical distribution function, denoted by  $F_{L_r}(x)$ , given by

$$\lim_{L_r \to \infty} F_{L_r}(x) \to F_r(x)$$
(23)

where

$$\frac{dF_r(x)}{dx} = f_r(x) = \begin{cases} \frac{1}{2\pi rx} \sqrt{(x - (1 - \sqrt{r})^2)((1 + \sqrt{r})^2 - x)} & (1 - \sqrt{r})^2 < x < (1 + \sqrt{r})^2 \\ 0 & otherwise \end{cases}$$
(24)

Using the above results, we obtain

$$\lim_{L_{r}\to\infty} \left( r \log_{2} d_{1} - \frac{1}{L_{r}} \sum_{i=1}^{L_{r}} \log_{2} d_{i} \right) = r \log_{2} \left[ L_{t} (1 + \sqrt{r})^{2} \right] - \lim_{L_{r}\to\infty} \left[ r \frac{1}{L_{t}} \sum_{i=1}^{L_{r}} \log_{2} \left( \frac{d_{i}}{L_{t}} \cdot L_{t} \right) \right]$$
$$= r \log_{2} (1 + \sqrt{r})^{2} - \lim_{L_{r}\to\infty} \left( r \frac{1}{L_{t}} \sum_{i=1}^{L_{t}} \log_{2} \frac{d_{i}}{L_{t}} \right)$$
$$= r \log_{2} (1 + \sqrt{r})^{2} - r \int_{(1 - \sqrt{r})^{2}}^{(1 + \sqrt{r})^{2}} \log_{2} x \cdot f_{r}(x) dx$$
$$= r \log_{2} (1 + \sqrt{r})^{2} - r \eta(r)$$
(25)

where  $\eta(r) = \int_{(1-\sqrt{r})^2}^{(1+\sqrt{r})^2} \log_2 x \cdot f_r(x) dx$ . We note that  $\eta(r) < \infty$ , for  $0 \le r < 1$ .

Finally, by substituting (25) into (21), we get

$$\lim_{L_r \to \infty} \frac{C_U - C_L}{L_r} = -(1 - r) \log_2(1 - r) + r \left( \log_2(1 + \sqrt{r})^2 - \eta(r) \right)$$
(26)

Clearly, when  $r \to 0$ , we have  $\lim_{L_r \to \infty} \frac{C_U - C_L}{L_r} \to 0$ , implying that when  $L_r \gg L_t$ , the capacity lower

and upper bounds converge to the same limit.

#### **D.** Optimal Antenna Selection

From the closed form expressions of  $C_L$  and  $C_U$  given in (15) and (20), respectively, we clearly distinguish the channel correlation effect on the capacity from the instantaneous channel effect given by  $\tilde{\mathbf{H}}_{w}$ . We can then proceed with antenna selection based on the LT-CS given by  $\mathbf{R}_{\Lambda_t}$  and  $\mathbf{R}_{\Lambda_r}$ . Clearly, the antenna selection criterion that maximizes  $C_L$  has to maximize the term  $\log_2\left(\prod_{i=1}^{L_r} q_t^{(i)}\right) + \log_2\left(\prod_{i=1}^{L_r} q_r^{(i)}\right)$  in (15). The resulting antenna set, thus, needs to be selected so as to maximize the determinant of  $\mathbf{R}_{\Lambda_t}$  and  $\mathbf{R}_{\Lambda_r}$  at the transmitter and receiver side, respectively.

Furthermore, a closer look at Eqn. (20) indicates that  $C_U$  is maximized according to the same selection criterion. It follows that the selected antenna subset that maximizes the upper bound also

maximizes the lower bound simultaneously. Moreover, according to Section III.C,  $C_L$  and  $C_U$  converge to the same limit when  $L_r \gg L_t$ . This implies that the antenna subset that maximizes both bounds also maximizes the capacity and is hence optimal.

So far, we have shown that to maximize the capacity, the antennas should be selected to maximize the determinant of the corresponding correlation matrix in both scenarios:  $L_r = L_t$  and  $L_r > L_t$ . Let

$$\xi_{\Lambda_{t}}^{C} = \prod_{i=1}^{L_{t}} q_{t}^{(i)} \text{ and } \xi_{\Lambda_{r}}^{C} = \prod_{i=1}^{L_{r}} q_{r}^{(i)}.$$
 (27)

Then, the capacity maximization criterion for joint transmit and receive antenna selection can be described as follows.

**Proposition 1:** For a given  $L_t$  and  $L_r$ , the optimal selected transmit antenna subset  $\Lambda_t^*$  and receive antenna subset  $\Lambda_r^*$  that maximize the capacity are given by

$$\Lambda_{\mathbf{r}}^* = \arg \max_{\Lambda_{\mathbf{r}}} \xi_{\Lambda_{\mathbf{r}}}^C$$
, and  $\Lambda_{\mathbf{t}}^* = \arg \max_{\Lambda_{\mathbf{t}}} \xi_{\Lambda_{\mathbf{t}}}^C$ 

where  $\xi_{\Lambda_r}^c$  and  $\xi_{\Lambda_t}^c$  are given by (27).

As mentioned before, the above derivation is based on the assumption that both  $\mathbf{R}_{\Lambda_t}$  and  $\mathbf{R}_{\Lambda_r}$ are of full rank (See Section II). Actually, this assumption can be relaxed. Using extensive simulations, we found that even when  $\mathbf{R}_{\Lambda_t}$  and  $\mathbf{R}_{\Lambda_r}$  are singular, the criterion is also applicable if we substitute  $L_t$  and  $L_r$  in (27) by  $rank(\mathbf{R}_{\Lambda_t})$  and  $rank(\mathbf{R}_{\Lambda_r})$ , respectively. In other words, when the number of antennas to be selected,  $L_t$ , is larger than  $rank(\mathbf{R}_{\Lambda_t})$ , the selected  $L_t$  set of antennas should include those who maximize  $\prod_{i=1}^{rank(\mathbf{R}_{\Lambda_t})} q_t^{(i)}$ . Selection at the receiver side is similar.

## **IV. Antenna Selection Algorithms for Correlated Channels**

We describe here a selection process according to Proposition 1, which we shall denote by *Correlated Selection Algorithm* (CSA). This algorithm consists of creating all possible  $C_M^{L_r}$  ( $C_N^{L_r}$ )

antennas sets  $\Lambda_t$  ( $\Lambda_r$ ) with  $L_t$  ( $L_r$ ) out of *M* transmit (*N* receive) antennas. The corresponding  $\xi_{\Lambda_t}^c$  ( $\xi_{\Lambda_r}^c$ ) are computed and the one with the best measure, as described in Proposition 1, is selected. For simplicity, we only take the example of transmit selection. The description of CSA for receive selection is similar.

## ALGORITHM I CORRELATED SELECTION ALGORITHM (CSA)

Given  $L_t$ , let  $K = C_M^L$ , and generate all possible  $\Lambda_t$ 's:  $\Lambda_t^{(1)}$ ,  $\Lambda_t^{(2)}$ , ...,  $\Lambda_t^{(K)}$ . *Initialization:*   $\xi^C = 0$ ,  $\Lambda_t^* = \mathbf{0}_{\mathbf{L}_t \times \mathbf{1}}$ , *Recursion:* For m = 1: KCompute  $\xi_{\Lambda_t^{(m)}}^C$  for each  $\Lambda_t^{(m)}$ . If  $\xi_{\Lambda_t^{(m)}}^C \ge \xi^C$  Then  $\Lambda_t^* = \Lambda_t^{(m)}$ ,  $\xi^C = \xi_{\Lambda_t^{(m)}}^C$ . End if. End loop Output  $\Lambda_t^*$ .

Clearly, transmit and receive antenna selection are decoupled with CSA. Therefore, only  $C_M^{L_1} + C_N^{L_r}$  comparisons are needed with CSA instead of  $C_M^{L_1} \times C_N^{L_r}$  with the full exhaustive search algorithm. The latter consists of considering all possible antenna sets at both ends simultaneously in couples,  $(\Lambda_t, \Lambda_r)$ . Nevertheless, at either communication end, CSA employs an exhaustive search for the optimal antenna set, which will incur prohibitive computational complexity for a large M (or N). For instance, assume that M = 20 and  $L_r = 8$ , then a total of  $C_M^{L_r} = 125970$  comparisons are needed, which is still considerably computationally high.

To further reduce the complexity, we propose to apply a sub-optimal sequential selection approach instead of the exhaustive one used in CSA. This is briefly described as follows in the case of transmit selection. We begin by considering all the *M* transmit antennas. Antenna selection is then performed in stages. In the *i*<sup>th</sup> stage, all possible antenna subsets  $\Lambda_t^{(j)}(i)$  are obtained by removing only one antenna each time from the original set  $\Lambda_t^*(i-1)$  of the previous  $(i-1)^{th}$  stage. For each  $\Lambda_t^{(j)}(i)$ , we compute its corresponding determinant given by

$$\xi_{\mathbf{A}_{\mathbf{t}}^{(j)}(i)}^{C} = \prod_{k=1}^{\left|\mathbf{A}_{\mathbf{t}}^{(j)}(i)\right|} q_{t}^{(k)} \,.$$
(28)

For ease of reference, we will refer to the above determinant as the metric of each antenna subset  $\Lambda_t^{(j)}(i)$ . Then, the optimal antenna subset with the highest metric,  $\Lambda_t^*(i)$ , is selected. All other sub-sets are then removed and the search will continue as described above until only  $L_i$  antennas are left. For instance, consider transmit selection with  $L_i = 2$  and M = 6, as shown in Fig. 1. We label the transmit antennas as 1, 2, ..., 6. In particular, the notation  $\{i, j, ..., k\}$  represents the selection of Antenna i, j, and k from the set of the M transmit antennas. It can be seen that four stages of search are required. In the 1<sup>st</sup> stage, 5 transmit antennas are chosen out of the already selected optimal antenna subset in the first stage. This process is continued until we obtain the best selected transmit antenna set with  $L_i = 2$  antennas. This sequential search algorithm shall be referred to as *low-complexity correlated selection algorithm* (L-CSA) and is described below in the case of transmit selection.

## ALGORITHM II LOW-COMPLEXITY CORRELATED SELECTION ALGORITHM (L-CSA)

Let  $\Lambda_t^*(i)$  denote the optimal selected antenna set at the *i*-th stage, and label the transmit antennas as 1, 2, ..., *M*. *Initialization:*  L = M, i = 0,  $\Lambda_t^*(0) = \{1, 2, ..., M\}$  *Recursion:* While  $(L > L_t)$ a) i = i + 1; b) For j = 1 : L  $\Lambda_t^{(j)}(i) = \Lambda_t^*(i-1) - \{j\}; \quad \xi_{\Lambda_t^{(j)}(i)}^C = \prod_{i=1}^{|\Lambda_t^{(j)}(i)|} q_t^{(i)}$ . End c)  $\Lambda_t^*(i) = \arg \max_{\Lambda_t^{(j)}(i)} \xi_{\Lambda_t^{(j)}(i)}^C; \quad L = L - 1$ . End While. Output  $\Lambda_t^*(i)$ . The optimal selected receive antenna set  $\Lambda_r^*$  can be obtained in a similar way. Clearly, L-CSA requires much fewer comparisons than CSA. Specifically, only  $\sum_{i=L_r+1}^M i + \sum_{j=L_r+1}^N j$  comparisons are

needed instead of  $C_M^{L_1} + C_N^{L_2}$  with CSA. To further demonstrate the amount of complexity reduction obtained with L-CSA, we provide in Table I the number of required comparisons to find the optimal antenna set using both CSA and L-CSA. For instance, to select 8 out of 20 transmit and 12 out of 30 receive antennas, CSA needs about  $8 \times 10^7$  comparisons, while with L-CSA only 561 comparisons are needed. Nevertheless, despite its low complexity, L-CSA may not provide the optimal set that maximizes the capacity. As shown in Fig. 1, two sets in the first stage have the same value of the largest metric. That is, the sets inside the solid and dashed line boxes. In this case, L-CSA chooses the first found set. However, note that a better capacity may be obtained if the other set is considered. In this paper, we consider the simplest form of L-CSA, which always selects the first found set that maximizes the metric as described in (28) at any given stage, to provide the lowest possible complexity.

## **V. Simulation Results and Discussions**

In this section, we present simulation results that validate the selection criterion derived previously. We also compare the performance of CSA with both ISA and the conventional system. Performance is evaluated in terms of capacity averaged over 50000 frames. We consider uplink transmission and adopt the correlated channel model described in [11,12]. Linear arrangement of the antenna array is assumed at both the receiver (base station side) and transmitter (mobile side) with the antenna separation being 4 and 1/2 wavelengths, respectively. We also assume the "broadside" case as defined in [11], and that the incoming waves are uniformly distributed in the angular spread  $\Delta_r(\Delta_t)$  [12].

### A. Theoretical Results Validation

To confirm the optimality of our criterion, we compare the CSA selection to the exhaustive search one. The latter is obtained by using Monte Carlo simulations to find the best  $L_r(L_r)$  antennas that should be selected on the transmitter (receiver) side. To do so, we compute the corresponding capacity

14

for each possible antenna subset  $\Lambda_t(\Lambda_r)$ , and choose the one with the best performance. Table II presents a sample of the comparison results for receive selection under different receive angular spreads  $\Delta_r$  and different number of selected receive antennas  $L_r$ . Correlation is assumed to be existent only at the receiver side. We note that in some cases, two different  $\Lambda_r^{(i)}$  and  $\Lambda_r^{(j)}$  could have  $\xi_{\Lambda_r^{(j)}}^C = \xi_{\Lambda_r^{(j)}}^C$ , such is the case with  $\Lambda_r^{(1)} = \{1, 2, 6\}$  and  $\Lambda_r^{(2)} = \{1, 5, 6\}$  when  $\Delta_r = 10^\circ$  and  $L_r = 3$ . In this particular case, both subsets are optimal in terms of capacity maximization and  $\Lambda_r^*$  is thus not unique. We evaluated their corresponding capacity using Monte Carlo simulations and found a difference of the order of 1%. As a result, both selections are assumed optimal and we list both of them in the fourth column of Table II. Similar results are also obtained for the transmitter side. However, we do not include them for space limitations. Based on these results, it is obvious that the CSA selection results coincide very well with the simulation results. Thus, we conclude that our selection criterion, as derived in Section III, is indeed optimal.

#### **B.** Selection Gain

We compare the capacity of our proposed CSA and the random selection algorithm which we refer to as RSA so as to see how much gain can be obtained with the optimal selection. In particular, with RSA, we randomly select  $L_r$  receive antennas or  $L_t$  transmit antennas for each channel realization. Fig. 2 shows the comparison results under different values of  $L_r$ ,  $L_t$ ,  $\Delta_r$  and  $\Delta_t$ . We first consider the capacity gain at the receiver side. Particularly, we assume that correlation only exists at the receiver and for a system of N = 6 and M = 2, we select 3 receive antennas randomly or according to our proposed CSA. From Fig. 2, it can be seen that the optimal receive antenna selection using CSA can achieve a gain of at least 1 bit/s/Hz for 90% outage capacity when  $\Delta_r = 35^\circ$  and SNR = 20dB. As for the transmitt antenna selection, we consider a system of M = N = 6 and assume that correlation only exists at the transmitter side. A gain of 3 bit/s/Hz at 90% outage capacity is observed with CSA when  $\Delta_r = 90^\circ$  and SNR = 20dB.

We further investigate in what follows the capacity gain achieved with CSA compared to RSA with different values of  $L_r$  and  $\Delta_r$ . To do so, we evaluate here such gain as

15

 $(C_{CSA} - C_{RSA})/C_{RSA} \times 100\%$ , where  $C_{CSA}$  and  $C_{RSA}$  denote the 90% outage capacity obtained by CSA and RSA, respectively. Fig. 3 provides the gain results for  $\Delta_r \in [5^\circ, 75^\circ]$  and  $L_r = 2, 3$  and 4. A closer observation of this figure indicates that CSA exhibits significant gain compared to RSA only for an angular spread, i.e.,  $\Delta_r \in [10^\circ, 60^\circ]$ . In fact, when  $\Delta_r$  is relatively small, i.e.,  $\Delta_r < 10^\circ$ , the correlation matrix tends to be singular with a unit rank. No matter what antennas are selected, there is only one independent dimension for the whole channel and thus the capacity gain obtained with CSA is very low. A similar observation is also noticed for a relatively large angular spread, i.e.,  $\Delta_r > 60^\circ$ . In this case, however, the correlation between different antennas tends to be zero implying that  $\mathbf{R}_r = \mathbf{I}_{N\times N}$ . Antenna selection does not, hence, affect the capacity. As a result, we can conclude that CSA presents a significant gain compared with RSA when the channel links are neither independent nor severely correlated. Besides, Fig. 3 indicates that such gain decreases with an increasing  $L_r$ . This is because the larger the antenna set to be selected, the higher the chances that RSA would select the same antenna set as CSA. Similar results can be obtained for transmit selection. However, we do not include them here for space limitation.

### C. Performance Comparison with Instantaneously Selected System

We compare here the capacity performance of CSA and ISA. Recall that antenna selection in ISA is performed according to the exact channel state information per channel instance. For every realization of the channel matrix  $\mathbf{H}$ , a complete set of all the possible matrices  $\hat{\mathbf{H}}$  is created by eliminating all possible permutations of  $N-L_r$  rows (and/or  $M-L_r$  columns) from the matrix. Capacity is then computed for each possible  $\hat{\mathbf{H}}$  and the antenna set corresponding to  $\hat{\mathbf{H}}$  that maximizes the capacity is selected.

In Fig. 4, we consider antenna selection at the transmitter side. Capacity cdf results of both CSA and ISA for  $L_t = 2$  and 4,  $\Delta_t = 60^\circ$  and  $120^\circ$  with an SNR of 20 dB are provided in this figure. The number of receive antennas is fixed to be 6, and correlation is assumed to only exist at the transmitter side. It can be seen that CSA can always achieve nearly the same capacity as ISA even for large angular spreads. Recall that our proposed CSA is based on LT-CS provided by the correlation matrices.

With an increasing angular spread which implies a more and more uncorrelated channel, CSA will incur some performance degradation compared with ISA since capacity in this case is affected by IC-SI rather than by LT-CS. Nevertheless, from Fig. 4 it can be observed that this degradation is rather slight. Even for an angular spread of  $120^{\circ}$ , 90% outage capacity of CSA is only 0.4 bit/s/Hz less than that of ISA with  $L_t$  =4. As a result, we conclude that in correlated channels, antenna selection can be based on LT-CS instead of IC-SI with very slight capacity loss and significant complexity reduction. Similar results are also obtained for the receiver side, but we do not present them here due to limited space.

### D. Performance Comparison with the Conventional System

We compare the capacity achieved with CSA and the conventional system without selection (i.e.  $L_t = M$  and  $L_r = N$ ) to highlight the importance of antenna selection in correlated channels. Fig. 5 shows the capacity cdf curves of CSA and the conventional system. It can be observed that using less receive antennas will lead to a decrease of capacity. The capacity cdf curve of CSA with  $L_t = 2$ ,  $L_r = 6$  is always on the right side of those of CSA with  $L_t = 2$ ,  $L_r = 3$  or 5. However, using less transmit antennas may actually increase the capacity in some high correlated scenario. As Fig. 5 shows, when  $\Delta_t = 60^\circ$  and  $\Delta_r = 30^\circ$ , the capacity of CSA with  $L_t = 2$ ,  $L_r = 6$  (or 5) is larger than that of the conventional system ( $L_t = 6$ ,  $L_r = 6$ )! While with increasing  $\Delta_t$  and  $\Delta_r$ , the conventional system can achieve more capacity than CSA. For instance, with a  $\Delta_t = 120^\circ$  and  $\Delta_r = 60^\circ$ , the conventional system can system gains about 4 bit/s/Hz more than CSA with  $L_t = 2$ ,  $L_r = 6$ .

To show the effect of transmit selection and receive selection more clearly, Fig. 6 and Fig. 7 present the capacity of receive selection and transmit selection, respectively. In Fig. 6, we assume that no correlation exists at the transmitter and the number of transmit antennas is fixed to be 6. Clearly, the conventional system always outperforms CSA with receive antenna selection. Moreover, the capacity loss increases with a decreasing  $L_r$ . This is explained by the fact that when less receive antennas are selected, the overall diversity order decreases and so does the system capacity. Such a loss is also observed to decrease with a decreasing  $\Delta_r$ . In fact, a closer observation of Fig. 6 indicates

that when  $\Delta_r = 30^\circ$  and  $L_r = 2$ , CSA presents around a 4 bits/s/Hz loss compared with the conventional system at 90% outage capacity. With the same selected number of receive antennas, only a loss of the order of 1 bit/s/Hz is observed when  $\Delta_r = 10^\circ$ . This is explained by the fact that when  $\Delta_r$  decreases, the channel links becomes severely correlated. The capacity of the conventional system does not hence present much improvement compared to CSA. As a result, we conclude that CSA with receive antenna selection presents capacity degradation compared with the conventional system. In other words, receive selection always introduces performance degradation. Such degradation will increase with a decrease in the number of selected receive antennas and an increase of angular spread.

On the other hand, using less transmit antennas may boost the capacity. In order to show the effect of  $L_t$ 's and angular spreads on capacity more clearly, we plot the curves of 90% outage capacity versus  $\Delta_t$  with different  $L_t$ 's. Similarly, here no correlation is assumed to be at the receiver and the number of receive antennas is fixed to be 6. As Fig. 7 shows, under a highly correlated channel, i.e.  $\Delta_t \leq 60^\circ$ , CSA with  $L_t = 2$  provides the highest capacity. As the angular spread increases, the capacity with  $L_t = 2$  will converge to a constant quickly, while the capacity with more selected transmit antennas still go on climbing. Finally, when  $\Delta_t$  increases to 180°, the conventional system ( $L_t = M = 6$ ) achieves the best capacity performance. This phenomenon is coincident with what observed in [19]. Actually, if the number of transmit antennas exceeds the rank of the channel matrix, capacity will decrease with an increase of  $L_t$ . The highest capacity is always achieved when the number of transmit antennas is equal to the rank of the channel matrix. From Fig. 7, it is clear that only for high values of  $\Delta_t$  (i.e.,  $\Delta_t \ge 160^\circ$ , which implies a nearly uncorrelated channel), the conventional system outperforms CSA. As a result, we conclude that in correlated channels, the conventional system does not always provide the best capacity performance. In particular, optimal transmit antenna selection can bring capacity gain.

So far, we have shown that the optimal receive antenna selection presents some capacity degradation compared with the conventional system. The optimal transmit antenna selection, on the other hand, may enhance the capacity. Therefore, in correlated channels CSA with a proper  $L_t$  and

 $L_r$  can actually achieve more capacity than the conventional system. Proper antenna selection can not only be used to decrease the number of RF chains, but also as an effective means to improve the performance.

#### E. Performance Comparison of CSA with L-CSA

In Section IV, it has been shown that with L-CSA the complexity can be reduced dramatically. We further present the performance comparison of L-CSA and CSA in the case of receive selection. Their 90% outage capacity versus  $\Delta_r$  curves for different values of  $L_r$  are plotted in Fig. 8. It can be seen that L-CSA provides nearly the same capacity as CSA. In particular, it was found that the selection results using L-CSA are usually the same as those using CSA. As a result, we conclude that L-CSA can perform closely to CSA but with much lower complexity.

## VI. Conclusions

In this paper, we derived the capacity maximization criterion for transmit and receive antenna selection according to LT-CS. We showed that in correlated channels, our algorithm, which we refer to as the CSA scheme, can achieve nearly the same capacity as the ISA scheme while dramatically decreasing the complexity since only the correlation matrix is needed for selection instead of the instantaneous channel state information. It was also shown that with optimal transmit antenna selection, CSA can even achieve performance gain over the conventional system. Finally, we proposed a low-complexity selection algorithm L-CSA which can achieve very close performance to CSA but with much lower complexity.

## Appendix I. Proof of Lemma 1

Assume that **X** and **Y** are both  $n \times n$  non-negative definite Hermition matrices. Then, from [18] we know that

$$\lambda_{XYX}^{(i)} \le \lambda_{Y}^{(1)} \lambda_{x^{2}}^{(i)}, \quad i = 1, ..., n.$$
 (I.1)

Applying (I.1) with  $\mathbf{X} = \mathbf{B}^{1/2}$  and  $\mathbf{Y} = \mathbf{A}^* \mathbf{A}$ , we have

$$\lambda_{\mathbf{B}^{1/2}\mathbf{A}^*\mathbf{A}\mathbf{B}^{1/2}}^{(i)} \le \lambda_{\mathbf{A}^*\mathbf{A}}^{(1)}\lambda_{\mathbf{B}}^{(i)}, \quad i = 1, ..., n.$$
(I.2)

Furthermore, using the fact that  $\lambda_{AA^*}^{(i)} = \lambda_{A^*A}^{(i)}$ , i = 1, ..., m, we know that

$$\lambda_{ABA^*}^{(i)} = \lambda_{B^{1/2}A^*AB^{1/2}}^{(i)}, \quad i = 1, ..., m.$$
(I.3)

As a result, we get

$$\lambda_{ABA^*}^{(i)} \le \lambda_{A^*A}^{(i)} \lambda_B^{(i)}, \quad i = 1, ..., m.$$

### REFERENCES

[1] R. D. Murch and K. B. Letaief, "Antenna systems for broadband wireless access," *IEEE Communications Magazine*, Vol. 40, No. 4, pp. 76-83, April 2002.

[2] K. K. Wong, R. D. Murch, and K. B. Letaief, "Performance enhancement of multiuser MIMO wireless communications systems," *IEEE Transactions on Communications*, Vol. 50, No. 12, pp. 1960-1970, Dec. 2002.

[3] P. W. Wolniansky, G. J. Foschini, G. D. Golden and R. A. Valenzuela, "V-BLAST: an architecture for realizing very high data rates over the rich-scattering wireless channel," in *Proc. ISSSE'98*, pp. 295-300, 1998.

[4] E. Telatar, "Capacity of multi-antenna Gaussian channels," AT&T Bell Labs Internal Tech. Memo., June 1995.

[5] M. Z. Win and J. H. Winters, "Analysis of hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Trans. Commun.*, Vol. 47, No. 12, p. 1773-1776, Dec. 1999.

[6] M. Z. Win and J. H. Winters, "Virtual branch analysis of symbol error probability for hybrid selection/maximal-ratio combining in Rayleigh fading," *IEEE Trans. Commun.*, Vol. 49, pp. 1926-1034, Nov. 2001.

[7] R. W. Heath Jr. and A. Paulraj, "Antenna selection for spatial multiplexing systems based on minimum error rate," in *Proc. ICC'01*, pp. 2276-2280, June, 2001.

[8] S. Sandhu, R. U. Nabar, D. A. Gore, and A. Paulraj, "Near-optimal selection of transmit antennas for a MIMO channel based on Shannon capacity," in *Proc.* 34<sup>th</sup> Asilomar Conference on Signals, Systems and Computers, pp. 567-571, Oct. 2000.

[9] A. Molisch, M. Win and J. Winters, "Capacity of MIMO systems with antenna selection," in *Proc. ICC'01*, Helsinki, Finland, pp. 570-574, 2001.

[10] A. Gorokhov, D. Gore, and A. Paulraj, "Performance bounds for antenna selection in MIMO systems," in *Proc. ICC'03*, pp. 3021-3025, May 2003.

[11] D. S. Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn, "Fading correlation and its effect on the capacity of multi-element antenna systems," *IEEE Trans. Commun.*, Vol. 48, No. 3, pp. 502-513, 2000.

[12] D. Gesbert, H. Bolcskei, D. A. Gore, and A. J. Paulraj, "MIMO wireless channels: capacity and performance prediction," in *Proc. Globecom'00*, pp. 1083-1088, 2000.

[13] D. A. Gore, R. W. Heath, and A. J. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Commun. Letters*, Vol. 6, No. 11, pp. 491-493, Nov. 2002.

[14] D. Gore, R. Heath and A. Paulraj, "Statistical antenna selection for spatial multiplexing systems," in Proc.

ICC'02, New York, pp. 450-454, May 2002.

[15] D. A. Gore, and A. J. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Trans. Signal Processing*, Vol. 50, No. 10, pp. 2580-2588, Oct. 2002.

[16] J. W. Silverstein, "Eigenvalues and eigenvectors of large dimensional sample covariance matrices," *Contemporary Mathematics*, Vol. 50, pp. 153-159, 1986.

[17] C.-N Chuah, D. N. C. Tse, J. M. Kahn, and R. A. Valenzuela, "Capacity scaling in MIMO wireless systems under correlated fading," *IEEE Trans. Inf. Theory*, Vol. 48, No. 3, pp. 637-650, Mar. 2002.

[18] S. Wang, and Z. Jia, Inequalities in Matrix Theory, Anhui: Educational Publishing Company, 1994.

[19] D. A. Gore, and R. U. Nabar, and A. Paulraj, "Selecting an optimal set of transmit antennas for a low rank matrix channel," in *Proc. ICASSP'00*, pp. 2785-2788, 2000.

Algorithm	Complexity	<i>M</i> =6, <i>N</i> =6	M = 10, N = 15		<i>M</i> =20, <i>N</i> =30	
		$L_t = 3, L_r = 3$	$L_t = 2, L_r = 4$	$L_t = 5, L_r = 8$	$L_t = 3, L_r = 6$	$L_t = 8, L_r = 12$
CSA	$C_M^{L_t} + C_N^{L_r}$	40	1410	6687	594910	86619000
L-CSA	$\sum_{i=L_r+1}^M i + \sum_{j=L_r+1}^N j$	30	162	124	648	561

 TABLE I

 COMPLEXITY COMPARISON BETWEEN CSA AND L-CSA

TABLE II Receive Selection Results Using the Proposed Selection Scheme and Monte Carlo Simulation When N = 6 and M = 2

Angular Spread	Number of Selected	CSA selection	Monte Carlo Exhaustive	
Δ	Antennas L		Simulation	
-r	r		(SNR=20dB)	
	2	{1, 6}	{1, 6}	
100	3	{1, 2, 6} or {1, 5, 6}	$\{1, 2, 6\}$ or $\{1, 5, 6\}$	
10	4	$\{1, 2, 5, 6\}$	$\{1, 2, 5, 6\}$	
	2	{1, 6}	{1, 6}	
50°	3	{1, 3, 6} or {1, 4, 6}	{1, 3, 6} or {1, 4, 6}	
50	4	{1, 3, 5, 6} or {1, 2, 4, 6}	{1, 3, 5, 6} or {1, 2, 4, 6}	
	2	{2, 6}	{2, 6}	
700	3	{1, 2, 6} or {1, 5, 6}	{1, 2, 6} or {1, 5, 6}	
/0	4	{1, 2, 5, 6}	{1, 2, 5, 6}	



Fig. 1: Different stages of selection with L-CSA at the transmitter side for  $L_t = 2$  and M = 6. The bold box represents the selected set at a given stage.



Fig. 2: Capacity cdf curves of CSA and RSA for different values of  $L_r$ ,  $L_t$ ,  $\Delta_r$  and  $\Delta_t$  for SNR = 20dB



**Fig.3**: Capacity gain due to optimal receive selection vs.  $\Delta_r$  for different values of  $L_r$ . N = 6, M = 2, and SNR = 20 dB. No correlation exists at the transmitter.



**Fig. 4**: Capacity cdf curves of CSA and ISA with different values of  $L_t$  and  $\Delta_t$  when N = M = 6, and SNR = 20dB. No correlation exists at the receiver.



**Fig. 5**: Capacity cdf curves of CSA and the conventional system (with  $L_t = M = 6$ ,  $L_r = N = 6$ ) for different values of  $\Delta_t$ ,  $\Delta_r$ ,  $L_t$ , and  $L_r$ . SNR = 20dB. The capacity cdf curves of the conventional system are drawn in dash-dot lines.



Fig. 6: Capacity cdf curves of CSA for receive selection and the conventional system. N = M = 6, and SNR=20dB. No correlation exists at the transmitter. The capacity curves of the conventional system are drawn in dash-dot lines.



**Fig. 7**: 90% outage capacity vs.  $\Delta_t$  curves of CSA for transmit selection and the conventional system, for N = M = 6, and SNR=20dB. No correlation exists at the receiver. The curve of the conventional system is marked with "+".



Fig. 8: 90% outage capacity vs.  $\Delta_r$  curves of CSA and L-CSA for receive selection with different values of  $L_r$ , for N = 6, M = 2, SNR=20dB. No correlation exists at the transmitter.