Iterative Antenna Selection for Multi-Stream MIMO under a Holistic Power Model

Xingyu Zhou, *Student Member, IEEE,* Bo Bai, *Member, IEEE,* and Wei Chen, *Senior Member, IEEE*

***Abstract*—The design of wireless communication systems under holistic power models has been being a hot topic in green communications. In this letter, we shall investigate the antenna selection in multi-stream MIMO with circuit power consumption. The difficulty lies in the fact that the optimal active antennas can only be selected by exhaustive search. To reduce the complexity, two iterative properties of MIMO capacity with antenna selection under a holistic power model are first derived in this letter. An efficient iterative antenna selection algorithm can then be obtained directly from these properties. This algorithm enjoys a low complexity, and can be applied to both the transmitter and receiver. Simulation results will verify that the proposed algorithm achieves the near-optimal performance compared to exhaustive search.**

***Index Terms*—MIMO, antenna selection, multi-stream, iterative, holistic power model.**

I. Introduction

MIMO (multiple-input-multiple-output) is an emerging technology that significantly increases the data rates and reliability for communication systems [1]. Conventionly, more antennas achieve higher multiplexing and diversity gains. However, adopting multiple antennas will consume more circuit power when a holistic power model is considered [2]. Therefore, how to design energy-efficient schemes yet incur little performance loss in MIMO systems is highly important.

Antenna selection schemes in which only a subset of available antennas are active for transmitting or receiving could reduce the power consumption. In conventional antenna selection, the circuit power consumption is not considered and the constraint is only the transmission power. The goal is just to select an optimal subset of antennas to maximize the channel capacity. The authors in [3] proposed a fast antenna selection for point-to-point MIMO systems. Antenna selection for MU-MIMO systems has been investigated in [4]. However, the works in [5], [6] have shown that the antenna selection under a holistic power model is significantly different from conventional works. In this case, a dynamic change of active antennas must be applied to improve the spectral and energy efficiency. Jiang and Cimini in [5] consider this problem in a single data stream MIMO system where the transmission power and active antennas are jointly optimized to maximize the energy efficiency. In [6], a comparison is made between the energy efficiency of the transmit beamforming and the transmit antenna selection with a single RF chain. However, the antenna selection scheme for multi-stream MIMO systems under a holistic power model is more complicated as the received SNR of each stream must be optimized, which has not been studied in a systematic way.

In this paper, we propose a low complexity and near-optimal antenna selection algorithm for capacity maximization in a multi-stream MIMO system where the circuit power is considered. The iterative algorithm is based on the observation that there exists an iterative property of capacity with antenna selection. Therefore, we select one antenna which leads to the highest increment of capacity at each step. More importantly, the marginal benefit of adding one more antenna can be proved to diminishes with iteration by using the Cauchy's interlace theorem, which is leveraged to reduce the complexity even further. Moreover, the proposed algorithm can be applied to both the receive and transmit antenna selection. Simulation results show that it achieves near-optimal performance in both low and high SNR regimes.

II. System Model

Consider a point-to-point MIMO system with N_t transmit and N_r receive antennas. Assume that the channel experiences flat fading. The signal model of the considered MIMO system is given by

$$\mathbf{y} = \mathbf{Hs} + \mathbf{n}, \tag{1}$$

where $\mathbf{H}$ is the channel matrix whose $N_r \times N_t$ entries are i.i.d complex circular symmetric Gaussian random variables with zero-mean and unit variance. $\mathbf{s}$ and $\mathbf{y}$ represent the transmitted and received signals, respectively. $\mathbf{n}$ is the additive white Gaussian noise vector, each of whose elements is circularly symmetric complex random variable with zero mean and variance N_0. The channel state information (CSI) is only perfectly known at the receiver. Therefore the instantaneous capacity of the signal model in Eq. (1) is given by [1]

$$C = \begin{cases} C_{\mathrm{r}}(\mathbf{H}) = \log\det(\mathbf{I}_{N_t} + \frac{\rho}{N_t}\mathbf{H}^H\mathbf{H}); \text{receive selection} \\ C_{\mathrm{t}}(\mathbf{H}) = \log\det(\mathbf{I}_{N_r} + \frac{\rho}{N_t}\mathbf{H}\mathbf{H}^H); \text{transmit selection} \end{cases} \tag{2}$$

where $\rho = P_t/N_0$ is the common SNR at each receive antenna. P_t is the transmission power.

In [7], the power consumption of RF chains and other circuits in MIMO is captured by

$$P_c = L_t \cdot P_{ct} + L_r \cdot P_{cr} + P_{c0}, \tag{3}$$

Manuscript received October 26, 2013. The associate editor coordinating the review of this letter and approving it for publication was W. Choi.

This paper is partially supported by the National Basic Research Program of China (973 Program 2013CB336600 and 2012CB316000), NSFC Excellent Young Investigator Award No. 61322111, Chuanxin Funding, MoE new century talent program under No. NCET-12-0302, Beijing Nova Program No. Z121101002512051, and National Science and Technology Key Project No. 2013ZX03003006-005 and 2013ZX03003004-002.

The authors are with the Department of Electronic Engineering, Tsinghua National Lab on Information Science and Technology (TNList), Tsinghua University, Beijing 100084, China (e-mail: eebobai@tsinghua.edu.cn).

Digital Object Identifier 10.1109/WCL.2013.111713.130754

where L_t and L_r are the number of active transmit and receive RF chains, respectively. P_{ct} and P_{cr} stand for the power consumed by each transmit and receive RF chain. P_{c0} is the power of frequency synthesizer and other units of circuits. Thus, the overall power consumption of MIMO system can be obtained as

$$P = \frac{1}{\eta_{pa}} \cdot P_t + P_c, \tag{4}$$

where η_{pa} is the drain efficiency of the power amplifier.

III. Antenna Selection Algorithm

In this section, we propose a low complexity algorithm to achieve the near-optimal antenna selection performance. Our aim is to find the optimal number and subset of active receive or transmit antennas that achieves the maximum capacity under the constraint of a total power consumption P. The best submatrix $\mathbf{H}_\Omega$ can only be obtained by exhaustive search which is complexity prohibitive. To avoid this problem, we propose a low complexity iterative algorithm. The fundamental idea behind this algorithm has two key points: 1) judiciously select one receive or transmit antenna which leads to the highest increment of capacity at each step; 2) decide when to stop the iteration in advance.

The proposed algorithm relys on the following two facts. One is that there exists an iterative property of channel capacity of MIMO systems where the receive or transmit antenna selection is adopted, which is stated in Lemma 1 and 2. The other one is that the marginal benefit of adding one more antenna is proved to be in descending order by using the Cauchy's interlace theorem, which is stated in Theorem 1.

We denote by $\mathbf{H}_n$ the channel matrix after n steps of selection of antenna. At the $(n+1)$th step, if the s^*th row or column of $\mathbf{H}$ is selected, the new $(n+1)\times N_t$ or $N_r \times (n+1)$ channel matrix is denoted by $\mathbf{H}_{n+1}$. $\mathbf{h}_s$ is a column vector which stands for the transpose of the sth row for receive antenna selection and the sth column for transmit antenna selection.

Lemma 1: With receive antenna selection, the capacity of a MIMO system with circuit power consumption could be expressed by the following iterative equation.

$$C_{\mathrm{r}}(\mathbf{H}_{n+1}) = C_{\mathrm{r}}(\mathbf{H}_n) + D_{\mathrm{r}}(\mathbf{H}_n) + \log(1+\Delta_{\mathrm{r},s,n}) \tag{5}$$

where

$$D_{\mathrm{r}}(\mathbf{H}_n) \triangleq \log\det(\mathbf{I}_n - \frac{\eta_{pa}P_{cr}}{N_0 N_t}\mathbf{H}_n(\mathbf{I}_{N_t} + \frac{\rho_{\mathrm{r},n}}{N_t}\mathbf{H}_n^H\mathbf{H}_n)^{-1}\mathbf{H}_n^H) \tag{6}$$

$$\mathbf{T}_{\mathrm{r},n} \triangleq (\mathbf{I}_{N_t}\frac{N_t}{\rho_{\mathrm{r},n+1}} + \mathbf{H}_n^H\mathbf{H}_n)^{-1} \tag{7}$$

$$\Delta_{\mathrm{r},s,n} \triangleq \mathbf{h}_s^H\mathbf{T}_{\mathrm{r},n}\mathbf{h}_s \tag{8}$$

Proof: See Appendix A. $\rho_{\mathrm{r},n}$ is defined in Eq. (23). ■

Lemma 2: With transmit antenna selection, the capacity of a MIMO system with circuit power consumption could be expressed by the following iterative equation.

$$C_{\mathrm{t}}(\mathbf{H}_{n+1}) = C_{\mathrm{t}}(\mathbf{H}_n) + D_{\mathrm{t}}(\mathbf{H}_n) + \log(1+\Delta_{\mathrm{t},s,n}) \tag{9}$$

where

$$D_{\mathrm{t}}(\mathbf{H}_n) \triangleq \log\det(\mathbf{I}_n - m_n\mathbf{H}_n^H(\mathbf{I}_{N_r} + \frac{\rho_{\mathrm{t},n}}{n}\mathbf{H}_n\mathbf{H}_n^H)^{-1}\mathbf{H}_n) \tag{10}$$

$$m_n = \frac{\rho_{\mathrm{t},n}}{n} - \frac{\rho_{\mathrm{t},n+1}}{n+1} \tag{11}$$

$$\mathbf{T}_{\mathrm{t},n} \triangleq (\mathbf{I}_{N_r}\frac{n+1}{\rho_{\mathrm{t},n+1}} + \mathbf{H}_n\mathbf{H}_n^H)^{-1} \tag{12}$$

$$\Delta_{\mathrm{t},s,n} \triangleq \mathbf{h}_s^H\mathbf{T}_{\mathrm{t},n}\mathbf{h}_s \tag{13}$$

Proof: See Appendix B. $\rho_{\mathrm{t},n}$ is defined in Eq. (28). ■

Remark 1: The two lemmas help to decouple the effect of antenna selection in MIMO systems.

We choose α as an indicator of the receive or transmit antenna selection whose value is either 'r' or 't'. Then $D_\alpha(\mathbf{H}_n)$ represents the effect of the circuit power consumption. The contribution of adding one antenna is depicted by the term $\log(1+\Delta_{\alpha,s,n})$. It motivates us to find the s^*th row or column that brings the largest contribution at each step, which is equivalent to the following problem.

$$s^* = \arg\max_s \Delta_{\alpha,s,n}. \tag{14}$$

Theorem 1: Both $D_\alpha(\mathbf{H}_n)$ and $\Delta_{\alpha,s,n}$ decrease as n increases. $D_\alpha(\mathbf{H}_n)$ is always negative and $\Delta_{\alpha,s,n}$ keeps positive.

Proof: Here we take the receive antenna selection as an example. The same proof works for the transmit antenna selection. To verify the property of $D_{\mathrm{r}}(\mathbf{H}_n)$, let $\mathbf{H}_n = \mathbf{U}\Sigma\mathbf{V}^H$ be the SVD decomposition of $\mathbf{H}_n$, where $\mathbf{U}$ and $\mathbf{V}$ are unitary matrices. We write

$$\mathbf{B}_n = \mathbf{I}_n - a\mathbf{H}_n(\mathbf{I}_{N_t} + b_n\mathbf{H}_n^H\mathbf{H}_n)^{-1}\mathbf{H}_n^H, \tag{15}$$

where $a = \frac{\eta_{pa}P_{cr}}{N_0N_t}$, $b_n = \frac{\rho_{\mathrm{r},n}}{N_t}$ and $b_n - a = \rho_{\mathrm{r},n+1} \geq 0$. With the SVD decomposition and some linear algebra computations, we could obtain

$$\mathbf{B}_n = \mathbf{U}(1 - \frac{a\Sigma^2}{1+b_n\Sigma^2})\mathbf{U}^H. \tag{16}$$

Since $D_{\mathrm{r}}(\mathbf{H}_n) = \log\det(\mathbf{B}_n)$, it follows that

$$\begin{aligned} D_{\mathrm{r}}(\mathbf{H}_n) &= \log(\prod_{i=1}^{n}(1 - \frac{a\sigma_i^2}{1+b_n\sigma_i^2})) \\ &= \log(\prod_{i=1}^{n}(1 - \frac{a\lambda_i}{1+b_n\lambda_i})), \end{aligned} \tag{17}$$

where σ_i is the ith singular value of $\mathbf{H}_n$ and $\lambda_i = \sigma_i^2$ is the ith eigenvalue of $\mathbf{H}_n\mathbf{H}_n^H$. For every i, the term $(1 - \frac{a\lambda_i}{1+b_n\lambda_i})$ is less than 1. Thus $D_{\mathrm{r}}(\mathbf{H}_n)$ is always negative. Next, to deal with the descending property of $D_{\mathrm{r}}(\mathbf{H}_n)$, we note that

$$\mathbf{H}_{n+1}\mathbf{H}_{n+1}^H = \begin{bmatrix} \mathbf{H}_n\mathbf{H}_n^H & \mathbf{H}_n\mathbf{h}_{s^*} \\ \mathbf{h}_{s^*}^H\mathbf{H}_n^H & \mathbf{h}_{s^*}^H\mathbf{h}_{s^*} \end{bmatrix}, \tag{18}$$

where $\mathbf{h}_{s^*}^H$ is the selected row at the last step. Its ith eigenvalue is denoted by λ_i^*. As $b_{n+1} < b_n$, we have

$$\begin{aligned} D_{\mathrm{r}}(\mathbf{H}_{n+1}) &= \log(\prod_{i=1}^{n+1}(1 - \frac{a\lambda_i^*}{1+b_{n+1}\lambda_i^*})) \\ &< \log(\prod_{i=1}^{n}(1 - \frac{a\lambda_i^*}{1+b_n\lambda_i^*})\cdot(1 - \frac{a\lambda_{n+1}^*}{1+b_n\lambda_{n+1}^*})). \end{aligned} \tag{19}$$

The Cauchy's interlace theorem [8] states that the eigenvalues of a Hermitian matrix of order $n+1$ are interlaced with those of any principal submatrix of order n. Here, $\lambda_{n+1}^* \leq \lambda_n^* \leq \lambda_{n-1}^* \leq \cdots \leq \lambda_1^*$ lists the eigenvalues of $\mathbf{H}_{n+1}\mathbf{H}_{n+1}^H$. $\lambda_n \leq \lambda_{n-1} \leq \cdots \leq \lambda_1$ are the eigenvalues of $\mathbf{H}_n\mathbf{H}_n^H$. By using the Cauchy's interlace theorem, we have

$$\lambda_{n+1}^* \leq \lambda_n \leq \lambda_n^* \leq \cdots \lambda_2 \leq \lambda_2^* \leq \lambda_1 \leq \lambda_1^* \tag{20}$$

Define $f(\lambda) = 1 - \frac{a\lambda}{1+b_n\lambda}$ and it is easy to verify the function of λ is monotonically decreasing. So with (20), we can conclude from Eq. (17) and Eq. (19) that $D_{\mathrm{r}}(\mathbf{H}_{n+1}) < D_{\mathrm{r}}(\mathbf{H}_n)$.

Next, we shall verify the property of $\Delta_{\mathrm{r},s,n}$. Assume that at the nth step, the s^*th is selected. Thus we get $\mathbf{T}_{\mathrm{r},n+1} = (\mathbf{I}_{N_t}\frac{N_t}{\rho_{\mathrm{r},n+2}} + \mathbf{H}_n^H\mathbf{H}_n + \mathbf{h}_{s^*}\mathbf{h}_{s^*}^H)^{-1}$. Applying the Woodbury matrix formula to it yields

$$\mathbf{T}_{\mathrm{r},n+1} = \tilde{\mathbf{T}}_{\mathrm{r},n} - \tilde{\mathbf{T}}_{\mathrm{r},n}\mathbf{h}_{s^*}(1+\tilde{\Delta}_{\mathrm{r},s^*,n})^{-1}\mathbf{h}_{s^*}^H\tilde{\mathbf{T}}_{\mathrm{r},n}, \tag{21}$$

where $\tilde{\mathbf{T}}_{\mathrm{r},n} = (\mathbf{I}_{N_t}\frac{N_t}{\rho_{\mathrm{r},n+2}} + \mathbf{H}_n^H\mathbf{H}_n)^{-1}$, $\tilde{\Delta}_{\mathrm{r},s^*,n} = \mathbf{h}_{s^*}^H\tilde{\mathbf{T}}_{\mathrm{r},n}\mathbf{h}_{s^*}$. The SVD decompositions of $\mathbf{T}_{\mathrm{r},n+1}$, $\tilde{\mathbf{T}}_{\mathrm{r},n}$ and $\mathbf{T}_{\mathrm{r},n}$ can be denoted by $\mathbf{T}_{\mathrm{r},n+1} = \mathbf{V}\Sigma\mathbf{V}^H$, $\tilde{\mathbf{T}}_{\mathrm{r},n} = \mathbf{U}\Sigma_1\mathbf{U}^H$ and $\mathbf{T}_{\mathrm{r},n} = \mathbf{U}\Sigma_2\mathbf{U}^H$. As they are all symmetric matrices, the diagonal elements in Σ, Σ_1 and Σ_2 are all positive. Thus, $\Delta_{\mathrm{r},s,n}$ keeps positive. Since $\rho_{\mathrm{r},n+2} < \rho_{\mathrm{r},n+1}$, the elements in Σ_1 are component-wise less than that in Σ_2. Thus we have

$$\begin{aligned}\Delta_{\mathrm{r},s,n+1} &= \mathbf{h}_s^H\mathbf{T}_{\mathrm{r},n+1}\mathbf{h}_s \\ &= \mathbf{h}_s^H\tilde{\mathbf{T}}_{\mathrm{r},n}\mathbf{h}_s - \mathbf{h}_s^H\tilde{\mathbf{T}}_{\mathrm{r},n}\mathbf{h}_{s^*}(1+\tilde{\Delta}_{s^*,n})^{-1}\mathbf{h}_{s^*}^H\tilde{\mathbf{T}}_{\mathrm{r},n}\mathbf{h}_s \\ &< \mathbf{h}_s^H\mathbf{U}\Sigma_1\mathbf{U}^H\mathbf{h}_s \\ &< \mathbf{h}_s^H\mathbf{U}\Sigma_2\mathbf{U}^H\mathbf{h}_s = \mathbf{h}_s^H\mathbf{T}_{\mathrm{r},n}\mathbf{h}_s = \Delta_{\mathrm{r},s,n}.\end{aligned} \tag{22}$$

■

Theorem 1 reveals the fact that the marginal benefit of adding one more antenna is diminishing and even may be negative, which verifies that activating more antennas is not always the best choice under a holistic power model. This fact can be leveraged to reduce the complexity of computation even further. Once the capacity for the current iteration is less than the last one, we break from the loop and immediately obtain the optimal antenna selection.

Based on the lemmas and theorem stated before, an efficient iterative algorithm is proposed. The details of the algorithm are presented in Algorithm I. It can be seen that at each step the optimal selection is captured by Eq. (14). Once the capacity is decreased, stop the iteration immediately. Moreover, this algorithm involves both the receive and transmit antenna selection with the indicator α which stands for either 'r' or 't'. For example, if α has the value of 'r', the algorithm will implement the receive antenna selection and $\bar{\alpha}$ is equal to 't'. The maximum number of iterations is linear with $N_{\max}$ but it could be largely reduced by Theorem 1.

IV. Simulation Results

In this section, simulation results are provided to demonstrate the performance of the proposed algorithm. The value for the parameters P_{ct}, P_{cr} , P_{c0} and η_{pa} are 120mW, 85mW, 30mW, 0.35, which are adopted from [7]. d is the distance between the transmitter and receiver. Log-distance path loss with a exponent of 4 is adopted. N_t and N_r are both 8.

Algorithm I
The Antenna Selection Algorithm

Input: P, P_{cr}, P_{ct}, P_{c0}, N_0, $\mathbf{H}$, α
Output: $\mathbf{H}_\Omega$
Initial: $N_{\max} = (P - N_{\bar{\alpha}}P_{c\bar{\alpha}} - P_{c0})/P_{c\alpha}$,
capacity = zeros$(1, N_{\max})$, $\Theta = \{1, 2 \ldots N_\alpha\}$, $\mathbf{H}_0 = \emptyset$
for $s = 1, \ldots, N_\alpha$
  $\Delta_s := \|\mathbf{H}(s,:)\|^2$ (α = 'r') or $\Delta_s := \|\mathbf{H}(:,s)\|^2$ (α = 't')
end
for $n = 1, \ldots, N_{\max}$
  $s^* := \arg\max_{s\in\Theta}\Delta_s$
  $\Theta := \Theta - \{s^*\}$
  $\mathbf{H}_n := [\mathbf{H}_{n-1}; \mathbf{H}(s^*,:)]$ (α = 'r')
  or $\mathbf{H}_n := [\mathbf{H}_{n-1}, \mathbf{H}(:,s^*)]$ (α = 't')
  capacity(n) := compute$(\mathbf{H}_n, \rho_{\alpha,n})$
  if capacity(n) < capacity$(n-1)$
    break
  else if (α=='r')
    $\mathbf{T}_n := (\mathbf{I}_{N_t}\frac{N_t}{\rho_{\mathrm{r},n+1}} + \mathbf{H}_n^H\mathbf{H}_n)^{-1}$
    $\Delta_s := \mathbf{H}(s,:)\mathbf{T}_n\mathbf{H}(s,:)^H$, $s \in \Theta$
  else if (α=='t')
    $\mathbf{T}_n := (\mathbf{I}_{N_r}\frac{n+1}{\rho_{\mathrm{t},n+1}} + \mathbf{H}_n\mathbf{H}_n^H)^{-1}$
    $\Delta_s := \mathbf{H}(:,s)^H\mathbf{T}_n\mathbf{H}(:,s)$, $s \in \Theta$
  end
end
Return $\mathbf{H}_\Omega = \mathbf{H}_n$

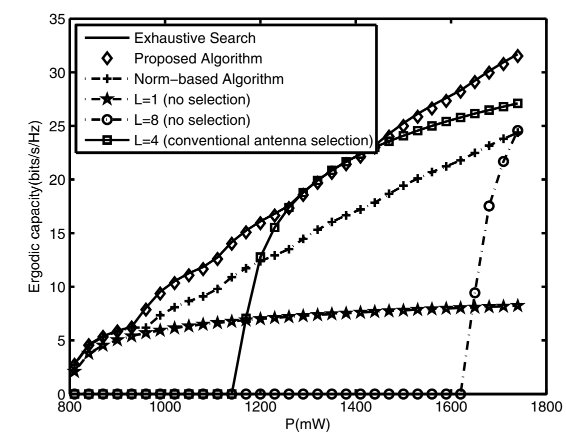

Fig. 1. Ergodic capacity VS. total power for $d = 600$m with receive selection

Figure 1 shows the ergodic capacity versus total power at the distance $d = 600$m. It is easy to find that near-optimal performance can be achieved by the proposed algorithm for different power consumption P. In addition, it can be seen that the proposed algorithm outperforms the conventional antenna selection algorithm and the norm-based algorithm in [5].

Figure 2 demonstrates the performance of different algorithms as a function of d for a given P. It can be seen that the proposed algorithm achieves near-optimal performance over all the transmission distances. Moreover, the gain over the norm-based algorithm is more significant for the moderate distance. It could be interpreted as follows. In low SNR regimes, the capacity is limited by the power. Therefore the

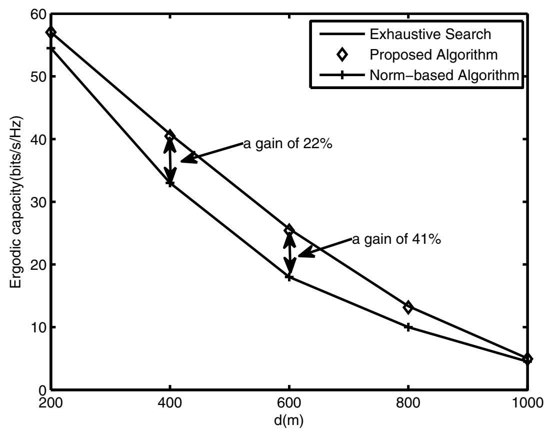

Fig. 2. Ergodic capacity VS. d with transmit selection.($P = 1500$mW)

row or column with the largest norm is the best choice. In the case of a large transmission distance, our proposed algorithm reduces to the norm-based algorithm and thus achieves the same performance. In high SNR regimes, the capacity is limited by the spatial degrees of freedom. Now activating more antennas is the best choice. Therefore as nearly all the antennas are active in the case of a large total power and a short distance, the gain over the norm-based algorithm isn't as significant as that for moderate distances.

V. Conclusion

An iterative algorithm has been developed to solve antenna selection problem for multi-stream MIMO systems where the circuit power consumption is considered. It can be adopted at both the transmitting and receiving ends. It greatly reduces the computational complexity yet with little performance loss when compared to the exhaustive search method. In the future work, we will extend the proposed algorithm to handle more practical cases like MU-MIMO systems and joint antennas selection.

Appendix A
Proof of Lemma 1

For the $(n + 1)$th step, the corresponding SNR can be obtained by

$$\rho_{\mathrm{r},n+1} = \frac{P_t^{(\mathrm{r},n+1)}}{N_0} = \frac{\eta_{pa}(P - P_{c0} - N_t P_{ct} - (n+1)P_{cr})}{N_0}, \tag{23}$$

where $P_t^{(\mathrm{r},n+1)} = \eta_{pa}(P - P_{c0} - N_t P_{ct} - (n+1)P_{cr})$. Thus, with Eq. (2) the capacity is as follows

$$C_\mathrm{r}(\mathbf{H}_{n+1}) = \log\det(\mathbf{I}_{N_t} + \frac{\rho_{\mathrm{r},n+1}}{N_t}\mathbf{H}_{n+1}^H\mathbf{H}_{n+1}). \tag{24}$$

Noting that

$$\mathbf{H}_{n+1}^H\mathbf{H}_{n+1} = \mathbf{H}_n^H\mathbf{H}_n + \mathbf{h}_s\mathbf{h}_s^H \tag{25}$$

and applying the matrix determinant lemma to Eq. (24), we obtain that

$$\begin{aligned} C_\mathrm{r}(\mathbf{H}_{n+1}) = \log\det(\mathbf{I}_{N_t} + \frac{\rho_{\mathrm{r},n+1}}{N_t}\mathbf{H}_n^H\mathbf{H}_n) + \\ \log(1 + \frac{\rho_{\mathrm{r},n+1}}{N_t}\mathbf{h}_s^H(\mathbf{I}_{N_t} + \frac{\rho_{\mathrm{r},n+1}}{N_t}\mathbf{H}_n^H\mathbf{H}_n)^{-1}\mathbf{h}_s) \end{aligned} \tag{26}$$

The first item on the right-side of the Eq. (26) can be expressed as Eq. (27) by using the generalization of matrix determinant lemma

$$\begin{aligned} \log\det(\mathbf{I}_{N_t} + \frac{\rho_{\mathrm{r},n+1}}{N_t}\mathbf{H}_n^H\mathbf{H}_n) = C_\mathrm{r}(\mathbf{H}_n) + \\ \log\det(\mathbf{I}_n - \frac{\eta_{pa}P_{cr}}{N_0N_t}\mathbf{H}_n(\mathbf{I}_{N_t} + \frac{\rho_{\mathrm{r},n}}{N_t}\mathbf{H}_n^H\mathbf{H}_n)^{-1}\mathbf{H}_n^H) \end{aligned} \tag{27}$$

We denote $\log\det(\mathbf{I}_n - \frac{\eta_{pa}P_{cr}}{N_0N_t}\mathbf{H}_n(\mathbf{I}_{N_t} + \frac{\rho_{\mathrm{r},n}}{N_t}\mathbf{H}_n^H\mathbf{H}_n)^{-1}\mathbf{H}_n^H)$ by $D_\mathrm{r}(\mathbf{H}_n)$. For the second item on the right-side of the Eq. (26), we denote $(\mathbf{I}_{N_t}\frac{N_t}{\rho_{\mathrm{r},n+1}} + \mathbf{H}_n^H\mathbf{H}_n)^{-1}$ by $\mathbf{T}_{\mathrm{r},n}$.

It's worth noting that the development of the proof could be degraded to that in [3] when the circuit power consumption is not taken into account.

Appendix B
Proof of Lemma 2

For the transmit antenna selection, the corresponding SNR of the $(n+1)$th step can be obtained by

$$\rho_{\mathrm{t},n+1} = \frac{P_t^{(\mathrm{t},n+1)}}{N_0} = \frac{\eta_{pa}(P - P_{c0} - N_r P_{cr} - (n+1)P_{ct})}{N_0}, \tag{28}$$

where $P_t^{(\mathrm{t},n+1)} = \eta_{pa}(P - P_{c0} - N_r P_{cr} - (n+1)P_{ct})$. With Eq. (2), the capacity of transmit antenna selection is as follows

$$C_\mathrm{t}(\mathbf{H}_{n+1}) = \log\det(\mathbf{I}_{N_r} + \frac{\rho_{\mathrm{t},n+1}}{n+1}\mathbf{H}_{n+1}\mathbf{H}_{n+1}^H) \tag{29}$$

we may adopt the similar methods of Eq. (26) and Eq. (27), obtain that

$$\begin{aligned} C_\mathrm{t}(\mathbf{H}_{n+1}) &= C_\mathrm{t}(\mathbf{H}_n) \\ &+ \log\det(\mathbf{I}_n - m_n\mathbf{H}_n^H(\mathbf{I}_{N_r} + \frac{\rho_{\mathrm{t},n}}{n}\mathbf{H}_n\mathbf{H}_n^H)^{-1}\mathbf{H}_n) \\ &+ \log(1 + \frac{\rho_{\mathrm{t},n+1}}{n+1}\mathbf{h}_s^H(\mathbf{I}_{N_r} + \frac{\rho_{\mathrm{t},n+1}}{n+1}\mathbf{H}_n\mathbf{H}_n^H)^{-1}\mathbf{h}_s), \end{aligned} \tag{30}$$

where $m_n = \frac{\rho_{\mathrm{t},n}}{n} - \frac{\rho_{\mathrm{t},n+1}}{n+1}$.

Using $\mathbf{T}_{\mathrm{t},n}$ and $D_\mathrm{t}(\mathbf{H}_n)$ for simplification, the conclusion of the lemma is verified. The selection is fedback to the transmitter through a noiseless feedback channel.

References

[1] E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecommun.*, vol. 10, no. 6, pp. 585–595, 1999.
[2] G. Li, Z. Xu, C. Xiong, C. Yang, S. Zhang, Y. Chen, and S. Xu, "Energy-efficient wireless communications: tutorial, survey, and open issues," *IEEE Wireless Commun. Mag*, vol. 18, no. 6, pp. 28–35, 2011.
[3] M. Gharavi-Alkhansari and A. B. Gershman, "Fast antenna subset selection in MIMO systems," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 339–347, 2004.
[4] M. Sadek, A. Tarighat, and A. H. Sayed, "Active antenna selection in multiuser MIMO communications," *IEEE Trans. Signal Process.*, vol. 55, no. 4, pp. 1498–1510, 2007.
[5] C. Jiang and L. Cimini, "Antenna selection for energy-efficient MIMO transmission," *IEEE Wireless Commun. Lett.*, vol. 1, no. 6, pp. 577–580, 2012.
[6] M. Kakitani, G. Brante, R. Souza, and M. Imran, "Energy efficiency of transmit diversity systems under a realistic power consumption model," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 119–122, 2013.
[7] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, 2005.
[8] S.-G. Hwang, "Cauchy's interlace theorem for eigenvalues of hermitian matrices," *American Mathematical Monthly*, pp. 157–159, 2004.